THE LIBRARY OF CONGRESS
1800

LC Card #

15-22266

HANDBOOK

OF

MILITARY INFRARED TECHNOLOGY

Editor

William L. Wolfe

The University of Michigan

1965

Office of Naval Research
Department of the Navy
Washington, D.C.

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402 - Price $3.75

UG487
W6

Handbook of Military Infrared Technology

* * *

Compiled by The University of Michigan under Contract Nonr 1224 (12) with the Office of Naval Research and ARPA Order 161 with the Advanced Research Projects Agency. Published by the Naval Research Laboratory and printed by the U.S. Government Printing Office. Reproduction is permitted for any purposes of the U.S. Government.

Library of Congress Catalog Card No: 65-62266

4 - FEB - 1
Copy ____ 1967

Contributors

James L. Alward, Research Associate, The University of Michigan, Institute of Science and Technology, Ann Arbor, Michigan. *Spatial Frequency Filtering*

Richard Barakat, Research Laboratory, Itek Corporation, Lexington, Massachusetts. *Optical Frequency-Response Techniques*

C. E. Dunning, Deputy Manager, Research and Development, Astrionics Division, Aerojet-General Corporation, Azusa, California. *Targets*

R. J. Hembach, Head, Thermal Testing, Grumman Aircraft Engineering Corporation, Bethpage, New York. *Spacecraft Thermal Design*

L. H. Hemmerdinger, Group Leader, O.A.O. Thermal Control, Grumman Aircraft Engineering Corporation, Bethpage, New York. *Spacecraft Thermal Design*

J. A. Jenney, Graduate Research Assistant, The University of Michigan, Institute of Science and Technology, Ann Arbor, Michigan. *Physical Constants and Conversion Factors*

Richard Kauth, Research Associate, The University of Michigan, Institute of Science and Technology, Ann Arbor, Michigan. *Backgrounds*

Thomas Limperis, Associate Research Physicist, The University of Michigan, Institute of Science and Technology, Ann Arbor, Michigan. *Detectors*

Lawrence D. Lorah, Senior Scientist, Mithras, Inc., Cambridge, Massachusetts. *Aerodynamic Influences on Infrared System Design*

K. Robert Morris, Graduate Research Assistant, The University of Michigan, Institute of Science and Technology, Ann Arbor, Michigan. (now with Conduction, Inc., Ann Arbor, Michigan). *Control Systems*

Fred E. Nicodemus, Senior Scientist, Sylvania Electric Systems, Electronic Defense Laboratories, Mountain View, California. *Radiation Theory; Targets*

Richard A. Phillips, Assistant in Research, The University of Michigan, Institute of Science and Technology, Ann Arbor, Michigan. *Physical Constants and Conversion Factors*

Gilbert N. Plass, Professor of Atmospheric and Space Sciences, Southwest Center for Advanced Studies, Dallas, Texas. *Atmospheric Phenomena*

Eugene S. Rubin, President, Mithras, Inc., Cambridge, Massachusetts. *Aerodynamic Influences on Infrared System Design*

Sol Shapiro, Department Head, General Precision, Inc., Wayne, New Jersey. *System Design*

David Silvermetz, Staff Engineer, Servo Corporation of America, Hicksville, New York. *Preamplifiers and Associated Circuits*

Warren J. Smith, Vice President – Research, Optical Systems Division, Infrared Industries, Inc., Santa Barbara, California. *Optics; Optical Systems*

Allan L. Sorem, Research Associate, Research Laboratories, Eastman Kodak Company, Rochester, New York. *Film*

Gwynn H. Suits, Head, The University of Michigan, Institute of Science and Technology, Ann Arbor, Michigan. *Film*

William L. Wolfe, Research Engineer, The University of Michigan, Institute of Science and Technology, Ann Arbor, Michigan

Arthur E. Woodward, Engineering Manager, Servo Corporation of America, Hicksville, New York. *Preamplifiers and Associated Circuits*

Harold Yates, Chief Engineer, Field Research and Systems Department, Barnes Engineering Company, Stamford, Connecticut. *Atmospheric Phenomena*

Preface

Origins

Infrared techniques, particularly as applied to military problems, have been developing in a gradual way for many, many years. One of the first main uses was in World War II in such applications as the DOVE guided bomb and the German *Lichtsprecher*. Since then, the Sidewinders, the Midas satellites, communication devices, and a myriad of other instruments and systems have been proposed, designed, developed, and used. Whereas in the early 1950's a small group of researchers could convene and discuss most of the infrared problems that were then facing the nation, now such meetings encompass between five hundred and a thousand people. The techniques have changed, the components have improved, and the means of use for these components have become more sophisticated and more complicated.

Early in 1961, Al Canada, who was then at ARPA—the Advanced Research Projects Agency—and the editor of this Handbook had a series of conversations including discussions of ways to improve communications among infrared workers and of methods by which the talents of these workers could be used more effectively. Both of us were aware of the capabilities of organizations like IRIS (the Infrared Information Symposia) and IRIA (the Infrared Information and Analysis Center), but we felt that something different would be useful.

In the field of microwaves and radar, one of the first unclassified treatments of their application to fairly complicated guidance and control and related problems was a series of books written by staff members of the Radiation Laboratory of Massachusetts Institute of Technology. Many of these became classics. We felt that the infrared field could benefit from a similar publication program. Unfortunately, monetary problems forced us to modify our first hopes drastically. The result is this Handbook.

ARPA, via the Office of Naval Research, contracted with the University of Michigan to supervise writing of this book, and ONR requested the Naval Research Laboratory to arrange for its publication. ARPA suggested that a technical writing firm be retained for assistance, and the University, after evaluating competitive bids, selected the McGraw-Hill Book Company, Training Materials and Information Services Division—now Information and Training Services Division of the F. W. Dodge Company, a Division of McGraw-Hill, Inc. This group, whose responsibility was to the University, selected the authors for the several chapters and sections; they were in turn responsible to this McGraw-Hill division.

The project was initiated in August of 1962. Much manuscript was prepared during the ensuing year. In June of 1963 The University of Michigan elected to complete the final editing and writing without the assistance of McGraw-Hill. Some chapters were then finally prepared, and all chapters were revised somewhat, edited, and copy edited. The first portion of the manuscript was sent to the Naval Research Laboratory on February 25 and the last in August of 1964.

Credits

Hundreds of people have helped. The people whose help should be acknowledged in this preface are those who have made the most substantial contributions to the Handbook, although this should in no way indicate that the less substantial contributions were not valuable. The list of contributors (page iii) gives the chapter authors with their present affiliations, and the titles of their chapters.

People at The University of Michigan who should receive particular credit for their endeavors on this Handbook are Mrs. Hilda Taft and the IRIA clerical staff. Mrs. Taft painstakingly copy edited virtually the entire manuscript and it was she who brought to my attention all the idiosyncracies, irregularities, and inconsistencies in symbols and equations. If they are incorrect it is because I failed to take proper action after they were brought to my attention. The IRIA clerical staff, Dorothy Curtis, Beatrice Godin, Myrtle Kreie, Beth Larson, Marie Nichols, and Sonya Kennedy kept track of all the correspondence and the many, many revisions and different versions of each of the chapters, as well as the stray figures and reference checking during this work. I am also indebted to Dr. George J. Zissis, who critically read a great deal of the manuscript and provided useful comments about many chapters but particularly about Chapters 2, 3, and 4. J. P. Livisay and John Duncan criticized Chapter 6, and John Duncan wrote one section of Chapter 6; Donald M. Szeles provided some additional material for Chapters 9 and 10; Roy J. Nichols criticized Chapter 21; John Gebhardt calculated considerable material for some of the tables in Chapter 22.

We received criticisms and comments from many workers not associated with The University of Michigan. Lucien M. Biberman, Institute for Defense Analyses, Washington, D.C., and Dr. E. D. McAlister, Director, Applied Oceanography Group, University of California, Scripps Institution of Oceanography, San Diego, California, provided criticisms for Chapter 5; Dr. Stanley S. Ballard, University of Florida, Department of Physics, Gainesville, Florida, provided useful comments about Chapters 7 and 8; Dr. R. Clark Jones, Polaroid Corporation, Cambridge, Massachusetts, Dr. Henry Levinstein, Physics Department, Syracuse University, Syracuse, New York, and W. L. Eisenman, Naval Ordnance Laboratory, Corona, California, criticized the tables of Chapter 11; Dr. J. Stanley Buller, Santa Barbara Research Center, Goleta, California, provided useful criticisms for Chapter 12; K. L. DeBrosse,

ITT Industrial Laboratories, Fort Wayne, Indiana, commented on Chapter 14; and T. P. Jackson, Aerojet-General Corporation, Azusa, California, provided useful criticisms for Chapter 20.

All these people have helped by their comments, criticisms, or creative writing. It is only fair and truthful to add that in the end I have taken it upon myself to make some changes either as a result of what was written, as a result of the criticisms, or in spite of both. I hope that my changes have improved the Handbook. If you find that you argue with or take exception to either what is in this book or what is not, you are arguing principally with my decisions.

The following individuals and companies provided information that was useful in preparing this Handbook:

J. G. Sample, *Raytheon Co.*
L. D. McGlauchlin, *Minneapolis-Honeywell Regulator Co.*
R. W. Weidmer, *Westinghouse Electric Corp.*
J. R. Kittler and P. L. Rice, *Garrett Corp.*
Allen Olsen, *Naval Ordnance Test Station*
Frank Kocsis, *Barnes Engineering Co.*
Joseph Jerger, *Servo Corporation of America*
Wayne McKusick, *Eastman Kodak Co.*
G. C. Higgins, *Eastman Kodak Co.*
John Howard, *Air Force Cambridge Research Laboratories*
Richard Rossi, *Malakar Laboratories*
J. Harkness, *Librascope Division, General Precision Co.*
John Harvell, *A. D. Little Co.*
L. L. Reynolds, *The RAND Corp.*
George R. Pruett, *Texas Instruments Incorporated*
Stanley Wallack, *Leesona Moos Corp.*
Radiation Electronics, Inc., *Infrared Detector Department*
E. M. Scott, *Engelhard Industries, Inc.*
Ralph Stair, *National Bureau of Standards*
Fish Schurman Corp.

WILLIAM L. WOLFE

Ann Arbor, Michigan
April 1965

Contents

Chapter		Page
	Contributors	iii
	Preface	v
1.	INTRODUCTION William L. Wolfe, *The University of Michigan*	1
2.	RADIATION THEORY William L. Wolfe, *The University of Michigan* Fred E. Nicodemus, *Sylvania Electronic Systems*	3
3.	ARTIFICIAL SOURCES	31
4.	TARGETS C. E. Dunning, *Aerojet-General Corporation* Fred E. Nicodemus, *Sylvania Electronic Systems*	57
5.	BACKGROUNDS Richard Kauth, *The University of Michigan*	95
6.	ATMOSPHERIC PHENOMENA Gilbert N. Plass, *Southwest Center for Advanced Studies* Harold Yates, *Barnes Engineering Company*	175
7.	OPTICAL COMPONENTS William L. Wolfe, *The University of Michigan*	281
8.	OPTICAL MATERIALS William L. Wolfe, *The University of Michigan*	315
9.	OPTICS Warren J. Smith, *Infrared Industries, Inc.*	369
10.	OPTICAL SYSTEMS Warren J. Smith, *Infrared Industries, Inc.*	421
11.	DETECTORS T. Limperis, *The University of Michigan*	457

12. DETECTOR COOLING SYSTEMS 519

13. FILM. 569
Allan L. Sorem, *Eastman Kodak Company*
Gwynn H. Suits, *The University of Michigan*

14. PREAMPLIFIERS AND ASSOCIATED CIRCUITS 583
Arthur E. Woodward and David Silvermetz,
Servo Corporation of America

15. OPTICAL FREQUENCY-RESPONSE TECHNIQUES 613
R. Barakat, *Itek Corporation*

16. SPATIAL FREQUENCY FILTERING 645
James Alward, *The University of Michigan*

17. CONTROL SYSTEMS. 661
K. R. Morris, *The University of Michigan*

18. SYSTEM DESIGN 729
Sol Shapiro, *General Precision, Inc.*
William L. Wolfe, *The University of Michigan*

19. INFRARED MEASURING INSTRUMENTS 757

20. SPACECRAFT THERMAL DESIGN 783
L. H. Hemmerdinger and R. J. Hembach,
Grumman Aircraft Engineering Corporation

21. AERODYNAMIC INFLUENCES ON
INFRARED SYSTEM DESIGN 825
Lawrence D. Lorah and Eugene Rubin, *Mithras, Inc.*

22. PHYSICAL CONSTANTS AND
CONVERSION FACTORS. 853
J. A. Jenney and Richard Phillips,
The University of Michigan

SUBJECT INDEX. 883

Special Figure, Table, and Text Acknowledgments

The parts of the *Handbook* itemized below are printed with the permissions indicated.

Chapter 2

Fig. 2-1: General Electric Company
Fig. 2-2: Block Associates, Inc.
Fig. 2-3: Admiralty Research Laboratory (Ministry of Defence, U.K.)

Chapter 3

Fig. 3-1: *Revue D'Optique*
Quotation in Section 3.1.3: American Society of Mechanical Engineers
Quotation in Section 3.2; Fig. 3-3, Fig. 3-4, and Fig. 3-5: National Bureau of Standards
Fig. 3-6: *Review of Scientific Instruments*

Chapter 4

Fig. 4-3 and 4-6: John Wiley & Sons, Inc.
Fig. 4-7: General Dynamics/Convair
Fig. 4-8 and 4-9: *Space/Aeronautics*
Fig. 4-11 and 4-12: Harvard University Press

Chapter 5

Fig. 5-1 to 5-9 incl.: Optical Society of America
Fig. 5-11, 5-12, and 5-18: Pergamon Press, Inc.
Fig. 5-13: Academic Press, Inc.
Fig. 5-14 and 5-15: *The Territorial Magazine*
Fig. 5-16 and 5-20: John Wiley & Sons, Inc.
Fig. 5-17: Clarendon Press, Oxford
Fig. 5-19: *Nature*
Fig. 5-21 and 5-22, also Tables 5-2 to 5-7 incl.: Aerojet-General Corporation
Fig. 5-23, 5-24, 5-25, also Table 5-8: The RAND Corporation
Fig. 5-33 to 5-38 incl.: McGraw-Hill, Inc.
Fig. 5-43 to 5-58 incl.: Air Weather Service (MATS)
Fig. 5-61 to 5-68 incl., 5-92, and 5-97: Optical Society of America
Fig. 5-93 and 5-94: Institute of Geophysics, University of California
Fig. 5-95 and 5-96: *Science*

Chapter 7

Fig. 7-1 to 7-8 incl.: Servo Corporation of America
Fig. 7-9 and 7-10: Barnes Engineering Co.
Fig. 7-12 and 7-13: U.S. Army Frankford Arsenal
Fig. 7-15 and 7-27: Optical Coating Laboratory, Inc.
Fig. 7-18: Barnes Engineering Company; *The Physical Review*
Fig. 7-19, 7-20, and 7-21: A. Smakula
Fig. 7-22 and 7-23: W. H. Freeman and Co.
Fig. 7-24: Bausch & Lomb, Inc.

Chapter 7 (continued)

Fig. 7-25: Eastman Kodak Co.
Fig. 7-26: Infrared Industries, Inc.
Fig. 7-28: Optical Society of America
Fig. 7-31, 7-32, and 7-34: McGraw-Hill, Inc.
Fig. 7-35, 7-36, and 7-37: Reinhold Publishing Corp.

Chapter 8

Fig. 8-19, 8-31, and 8-32, also Table 8-21: Optical Society of America
Fig. 8-21, 8-23 to 8-30 incl., 8-33, 8-34, and 8-35: Bausch & Lomb, Inc.
Fig. 8-22: Union Carbide Corp., Stellite Division
Fig. 8-37: Purdue Research Foundation
Fig. 8-39: Infrared Division, Research Dept., U.S. Naval Ordnance Laboratory, Corona, Calif.

Chapter 11

Section 11.4: Syracuse University and U.S. Naval Ordnance Laboratory, Corona, Calif.

Chapter 12

Table 12-1: Cryogenic Engineering Co.
Fig. 12-5: Barnes Engineering Co.
Fig. 12-7: AiResearch Manufacturing Co.
Fig. 12-9: Santa Barbara Research Center
Fig. 12-10: ITT Industrial Laboratories Division
Fig. 12-13: AiResearch Manufacturing Co.
Fig. 12-19: John Wiley & Sons, Inc.
Fig. 12-22 to 12-25 incl.: Borg-Warner Thermoelectrics

Chapter 13

Fig. 13-1 to 13-18 incl.: Eastman Kodak Co.

Chapter 17

Fig. 17-9 and 17-12, also Tables 17-1 and 17-2: McGraw-Hill, Inc.

Chapter 19

Fig. 19-10: Unicam Instruments Ltd.
Fig. 19-11, 19-12, and 19-16: Perkin-Elmer Corp.
Fig. 19-13: Barnes Engineering Co.
Fig. 19-14: Block Engineering, Inc.
Fig. 19-15: Beckman Instruments, Inc.
Fig. 19-17 and 19-19: W. H. Freeman & Co.
Fig. 19-20: Reinhold Publishing Corp.
Fig. 19-21: The Institute of Physics and The Physical Society, U.K.
Fig. 19-22: W. H. Freeman & Co.

Chapter 20

Fig. 20-1, 20-2, and 20-3: American Institute of Aeronautics and Astronautics
Fig. 20-4, 20-5, and 20-6: McGraw-Hill, Inc.
Fig. 20-7: American Meteorological Society
Table 20-3: John Wiley & Sons, Inc.
Fig. 20-12: American Institute of Aeronautics and Astronautics
Fig. 20-14 and 20-19: D. K. Edwards, K. E. Nelson, R. D. Roddick, and J. T. Gier

Chapter 21

Fig. 21-11 and 21-12: American Institute of Aeronautics and Astronautics

Chapter 22

Fig. 22-1, 22-4, and 22-5: Aerojet-General Corp.

Table 22-1: American Institute of Physics

Tables 22-6, 22-7, 22-8, 22-12, 22-14, and 22-18: Chemical Rubber Co. and John Wiley & Sons, Inc.

Table 22-9: University of Chicago Press

Tables 22-15 and 22-17: Chemical Rubber Co. and General Electric Co.

Handbook of Military Infrared Technology

Chapter 1
INTRODUCTION

William L. Wolfe

The University of Michigan

The chapters of this Handbook are arranged in a sequence that is now almost traditional, and it is logical. The radiators come first, then the medium of propagation, the receiver system, the transducers and electronics, and finally a number of special applications. Thus Chapters 2, 3, 4, and 5 deal with basic radiation laws, blackbody simulators, and the properties of targets and of backgrounds. Not very much attention is paid to the more difficult problem of calculating the amount and kind of gaseous radiation because it is not generally a problem for the systems engineer, whereas envelope calculations based on slide rules and formulas most certainly are. Traceability of instrument performance to the National Bureau of Standards is more and more a real question; therefore the entire problem of radiometrics has been dealt with in more detail than is usual for a handbook.

Chapter 6 deals with atmospheric absorption and contains some material on scattering and scintillation. Absorption processes and the calculation of absorption are relatively well known, and a detailed explanation of the theory and methods of calculation are given. The chief problem here is knowing the atmospheric composition. Much must still be done concerning the loss due to scattering, and with scintillation; here the terms are not even well defined.

The next group of chapters deals with optics and optical design. Considerable detail is given on design techniques because so little is available elsewhere. The basic nomenclature of Conrady is followed. Components and materials are discussed. A condensation of material contained in an IRIA* state-of-the-art report is given, with an augmentation on glasses. Although many optical components are bought by specification and fabrication to order, some and even some lenses do exist "on the shelf." As many as possible of these commercially available optical components are listed.

The chapter on detectors is relatively short, but the design engineer should find the extensive table of considerable use. For the first time a readily available useful display of most detector concepts appears in print. Methods of test vary, but those listed here have some measure of acceptance.

Detectors are often cooled, of course; Chapter 12 presents the most comprehensive table of coolers for infrared detectors ever published (more have probably come on the market since this book went to press). In addition, Stirling cycle systems and techniques using samples of solid hydrogen, helium, and the like are being developed, but they are not sufficiently well along for inclusion.

*Infrared Information and Analysis Center, a part of the Infrared Laboratories of the Institute of Science and Technology at The University of Michigan, Ann Arbor, Michigan. The report referred to is 2389-11-S, January 1959.

The chapters on film and preamplifiers conclude the treatment of the usual "parts" of an infrared system. Although there are many brochures on film and its performance, none has ever appeared that was couched in radiometric terms. It is apparently true that all infrared film is manufactured by the Eastman Kodak Company.

The remaining chapters deal with certain special features of infrared engineering. In this connection, the method of writing on systems design is of particular note. It should be self-evident that no two engineers do system design in the same fashion. Some general approaches have been discussed by such texts as Goode and Machol;* these usually include a respectable amount of probability mathematics and concepts, game theory, and block-diagram operations or signal-flow graphs. They also describe certain common-sense approaches to the simplification of the problem by dissection. To a large extent Chapter 18 is just such a description for infrared systems.

The authors hope that this Handbook will be used like most other handbooks. Occasionally the user will browse for ideas. More often he will be searching for the answer to a specific problem—the necessary data or the required formulas or techniques. The index should be the key to the answer for this need. It has been laboriously prepared to include references, cross references, and other helpful clues. The organization should serve those who browse and should help in specific searches by having, near the referenced, searched-for item, others that are closely associated with it.

This Handbook is not a state-of-the-art report on all phases of infrared systems and components; it is not and was not meant to be. But in some senses it has to be. A handbook is usually a source of useful information—data, equations, concepts, and techniques. It includes those things that are useful for undertaking certain development and research tasks, but it is not the last, up-to-the-minute word on all subjects.

As is so for every handbook, this one is neither completely up-to-date nor entirely comprehensive. The field is dynamic, and the sum of all material of interest to everyone in the field is an appreciable per cent of infinity. The references provide one clue for obtaining more information on any given subject. Another source is the various information centers dealing directly with infrared topics or touching upon infrared as a peripheral interest. Some of these centers are listed below. The Science Information Exchange, Suite 313, Universal Bldg., 1825 Connecticut Avenue, N.W., Washington, D.C. should be consulted for further lists of Centers in existence and their topics.

1-4. At the Institute of Science and Technology, The University of Michigan, Box 618, Ann Arbor, Michigan: IRIA, the Infrared Information and Analysis Center; BAMIRAC, the Ballistic Missile Radiation Analysis Center; TABSAC, the Target and Backgrounds Signature Analysis Center; BAC, the Background Analysis Center.

5. SPIA-LPIA, the Solid and Liquid Propellant Information Agency of the Applied Physics Laboratory of The Johns Hopkins University, 8621 Georgia Avenue, Silver Spring, Maryland.

6. RACIC, the Remote Areas Conflict Information Center, Battelle Memorial Institute, 505 King Avenue, Columbus, Ohio.

7. CINFAC, the Counterinsurgency Information Analysis Center, at American University, Washington, D.C.

8. IRIS, the Infrared Information Symposia, an organization devoted to appropriate timely dissemination of research and development results by meetings. Attendance is possible through Mr. Thomas B. Dowd, Office of Naval Research, 495 Summer Street, Boston, Massachusetts.

---

*H. Goode and R. Machol, *Systems Design, Control Systems Engineering,* McGraw-Hill Book Company, New York, 1957.

Chapter 2

RADIATION THEORY

William L. Wolfe

The University of Michigan

Fred. E. Nicodemus

Sylvania Electronic Systems

CONTENTS

2.1. Radiometric Quantities, Symbols, and Units 4
2.1.1. Radiometric Quantities as Field Concepts 7
2.1.2. Other Radiometric Quantities 7
2.2. Kirchhoff's Law: Emissivity and Blackbodies 9
2.3. Radiation Laws 9
2.3.1. Planck's Law 9
2.3.2. Quantum Rates in Blackbody Radiation 10
2.3.3. Stefan-Boltzmann Law 10
2.3.4. Rayleigh-Jeans and Wien Laws 10
2.3.5. The Wien Displacement Law 10
2.3.6. Maximum Difference Expression 10
2.4. Blackbody Slide Rules 11
2.4.1. The General Electric Rule 11
2.4.2. The Block Rule 11
2.4.3. The Admiralty Research Laboratory Rule 17
2.5. Blackbody Curves 17
2.6. Blackbody Tables 21
2.7. Radiation Geometry 22
2.7.1. Lambertian Sources 22
2.8. Distributed Radiators 23
2.9. Selective Radiators 23
2.10. Directional Reflectance and Emissivity 23
2.11. Summary of Equations and Constants 28

2. Radiation Theory

2.1. Radiometric Quantities, Symbols, and Units

The nomenclature, symbols, and units of the most important radiometric quantities are listed in Table 2-1, which is based on the recommendations of the Working Group on Infrared Backgrounds (WGIRB) [1,2]. They include and are consistent with American Standard Z58.1.1-1953.

TABLE 2-1. SYMBOLS, NAMES, AND UNITS OF RADIOMETRIC QUANTITIES

Symbol	*Name*	*Description*		*Units*
A	Area	Projected area		cm^2
Ω	Solid angle	–		sr
V	Volume	–		cm^3
U	Radiant energy	–		joule
u	Radiant energy density	Radiant energy per unit volume	$\frac{\partial U}{\partial V}$	joule cm^{-3}
P	Radiant power	Rate of transfer of radiant energy	$\frac{\partial U}{\partial t}$	w
W	Radiant emittance	Radiant power per unit area emitted from a surface	$\frac{\partial P}{\partial A}$	w cm^{-2}
H	Irradiance	Radiant power per unit area incident upon a surface	$\frac{\partial P}{\partial A}$	w cm^{-2}
J	Radiant intensity	Radiant power per unit solid angle from a point source	$\frac{\partial P}{\partial \Omega}$	w sr^{-1}
N	Radiance*	Radiant power per unit solid angle per unit projected area	$\frac{\partial^2 P}{\cos\theta\, \partial A\, \partial \Omega}$	w sr^{-1} cm^{-2}
P_λ	Spectral radiant power	Radiant power per unit wavelength interval	$\frac{\partial P}{\partial \lambda}$	w μ^{-1}
P_ν	Spectral radiant power	Radiant power per unit frequency interval	$\frac{\partial P}{\partial \nu}$	w sec

*Sometimes radiance is defined instead as the radiant power per unit area (not *projected* area) per unit solid angle $\left(N \equiv \frac{\partial^2 P}{\partial A\, \partial \Omega}\right)$. This is equally correct, but it is then necessary to insert the cos θ factor differently, *e.g.*, it is then $N \cos\theta$ which is invariant along a ray, and the radiance of a Lambertian surface varies with the cosine of the angle from the normal.

TABLE 2-1. SYMBOLS, NAMES, AND UNITS OF RADIOMETRIC QUANTITIES *(Continued)*

Symbol	*Name*	*Description*		*Units*
P_σ	Spectral radiant power	Radiant power per unit wave number interval	$\frac{\partial P}{\partial \sigma}$	w cm
W_λ	Spectral radiant emittance	Radiant emittance per unit wavelength interval	$\frac{\partial W}{\partial \lambda}$	w cm^{-2} μ^{-1}
H_λ	Spectral irradiance	Irradiance per unit wavelength interval	$\frac{\partial H}{\partial \lambda}$	w cm^{-2} μ^{-1}
J_λ	Spectral radiant intensity	Radiant intensity per unit wavelength interval	$\frac{\partial J}{\partial \lambda}$	w sr^{-1} μ^{-1}
N_λ	Spectral radiance	Radiance per unit wavelength interval	$\frac{\partial N}{\partial \lambda}$	w sr^{-1} cm^{-2} μ^{-1}

A second set of symbols and units (Table 2-2) is as written by Penner [3] and patterned after Worthing. In this system, a superscript 0, *e.g.*, R^0, indicates that the quantity is for a blackbody, and a subscript like ν or λ indicates partial differentiation. Thus, in this terminology $R_\lambda{}^0$ is the spectral radiancy of a blackbody in w cm^{-2} μ^{-1}; $R_\nu{}^0$ is the spectral radiancy of a blackbody in w cm^{-2} sec. Other investigators prefer to use the superscript or subscript *bb* or *b* to denote blackbody.

TABLE 2-2. PENNER RADIOMETRIC SYMBOLS

Symbol	*Definition*
$\mathfrak{E}$	Radiant energy
ρ	Radiant energy density
$\mathfrak{F}_e = \frac{\partial \mathfrak{E}}{\partial t}$	Power
$J = \frac{\partial \mathfrak{F}_e}{\partial \Omega}$	Radiant intensity
$R = \frac{\partial \mathfrak{F}_e}{\partial A_r}$	Radiancy: radiant flux per unit area from a source into a hemisphere
$W = \frac{\partial \mathfrak{F}}{\partial A_r}$	Radiant flux density (not of a source)
$B_{str} = \frac{1}{\cos\theta} \frac{\partial}{\partial A_r} \frac{\partial \mathfrak{F}_e}{\partial \Omega}$	Steradiancy or radiance
I or H or $R = \frac{\partial \mathfrak{F}_i}{\partial A_r}$	Irradiancy

Proposals have been made for altering the present names and symbols. The two now most in favor are those of R. Clark Jones and those of some NBS personnel. The latter would retain the words *radiant emittance, radiance, irradiance,* and *radiant intensity.* The system would, however, use the suffix *-ance* to describe measured properties of a particular sample and *-ivity* to indicate a property that is intrinsic with a material when used in connection with material rather than field quantities. Thus the reciprocal of the ratio of incident power (from a plane wave at normal incidence on a flat surface)

to the reflected power perpendicular to any sample is the normal reflectance. If the sample has optically flat polished surfaces the quantity is reflectivity. The obvious difficulty is in the words *emittance* and *emissivity,* although there are other difficulties.

Jones proposes a more general set of specifications. He deals with geometrical concepts. He calls the rate of flow of any quantity from a body per unit area *exitance,* the rate impinging per unit area *incidance,* the rate per unit area per unit solid angle *sterance,* and the rate per unit solid angle from a point source *intensity.* Then appropriate adjectival modifiers are used. Table 2-3 illustrates this system.

TABLE 2-3. JONES' PROPOSED TERMINOLOGY*

Geometric Quantity (Jones)	*Radiometric Quantity (Jones)*	*Radiometric Quantity (Standard)*
Emittance	Radiant exitance	Radiant emittance
Incidance	Radiant incidance	Irradiance
Sterance	Radiant sterance	Radiance
Intensity	Radiant intensity	Radiant intensity

*R. C. Jones, "Terminology in Photometry and Radiometry," *J. Opt. Soc. Am.*, 53, 11, 1314 (November 1963).

Another set of symbols (Table 2-4) is used in the thermal control industry and is gaining favor in a segment of the aerospace industries. It is also patterned after Worthing [4]. The systems are compared in Table 2-5.

TABLE 2-4. AEROSPACE RADIOMETRIC SYMBOLS

Symbol	*Term*	*Definition*
R	Radiancy	Power emitted per unit source area to a hemisphere
–	Steradiancy	Power emitted per unit area per unit solid angle.

TABLE 2-5. COMPARISON OF SYSTEMS OF TERMS*

WGIRB	*Penner*	*Aerospace*	*Jones*
Energy, U	Energy, $\mathfrak{E}$	Energy, Q	–
Energy density, u	Energy density, ρ	–	–
Power, P	Emitted power, $\mathfrak{F}_e$	Heat rate, q	–
Radiant emittance, W	Radiancy, R	Radiancy, R	Radiant emittance
Irradiance, H	Irradiancy, I or H or R	–	Radiant incidance
Radiant intensity, J	Radiant intensity, J	–	Radiant intensity
Radiance, N	Steradiancy or radiance, B_{str}	Steradiancy	Radiant sterance
	Radiant flux density (not source), W		

*NOTE ADDED IN PROOF: The Nomenclature Committee of the Optical Society of America has recently recommended that the flux density radiated from a source be called radiant exitance; the ratio of such flux from a sample to that of a blackbody be called emittance; and that of an opaque sample with perfect surfaces be called emissivity.

2.1.1. Radiometric Quantities as Field Concepts. The extension of the definitions of radiant emittance, radiant intensity, radiance, and irradiance to describe the properties of a radiant field, as well as the properties of a source, is of great utility. In particular, geometrically (neglecting attenuation by absorption, reflection, or scattering) the radiance at any point along a ray, in the direction of the ray, is invariant within an isotropic medium. In general, N/n^2 (where n is the index of refraction) is similarly invariant across a smooth boundary between two different media [5]. However, the determination of the radiometric properties of a source in an attenuating medium, from measurements made at a distance, always involves some assumptions about the nature of the attenuation, emission, and scattering of the intervening medium. Thus, source characteristics calculated from field quantities may be in error by an unknown amount. If no attempt is made in the calculation of the source characteristics to include the effects of the intervening medium, then the calculated quantities should have the adjectival modifier *apparent,* or they should indicate that the radiant field at the point of measurement is described.

2.1.2. Other Radiometric Quantities. Radiant absorptance, α, radiant reflectance, ρ, radiant transmittance, τ, and radiant emittance, ϵ, can be defined as:

$$\left.\begin{aligned} \alpha &= \frac{R_{absorbed}}{R_{incident}} \\ \rho &= \frac{R_{reflected}}{R_{incident}} \\ \tau &= \frac{R_{transmitted}}{R_{incident}} \\ \epsilon &= \frac{R_{emitted}}{R_{blackbody}} \end{aligned}\right\} \tag{2-1}$$

where R is the appropriate radiant quantity J, W, H, or N.

Table 2-6 provides a list of definitions for the "other" radiometric properties. The symbolism is still in a transitory stage. The symbols λ, T, H, N, etc., shown as subscripts (ϵ_{th} and α_s are also found elsewhere) do not indicate differentiation, *e.g.*, $\epsilon_\lambda \neq \partial\epsilon/\partial\lambda$, and for this reason WGIRB has recommended the use of $\epsilon(\lambda)$ rather than ϵ_λ for spectral emissivity, etc.

Radiant absorptance should not be confused with *absorption coefficient,* which is often represented by the symbol α; however, the symbol a for absorption coefficient is preferred.

The processes of absorption, reflection (including scattering), and transmission account for all incident radiation in any particular situation, and the absorptance, reflectance, and transmittance must add up to one:

$$\alpha + \rho + \tau = 1 \tag{2-2}$$

If a material is so opaque that it transmits no radiation, $\tau = 0$ and

$$\alpha + \rho = 1 \tag{2-3}$$

No specification has been made as to whether specular or diffuse quantities are indicated, and therefore the relations hold if the same specification is made for each of the quantities.

TABLE 2-6. "OTHER" RADIOMETRIC QUANTITIES*

ϵ	Emissance	The ratio of the rate of radiant energy emission from a body, as a consequence of its temperature only, to the corresponding rate of emission from a blackbody at the same temperature.
α	Absorptance	The ratio of the radiant energy absorbed by a body to that incident upon it.
ρ	Reflectance	The ratio of the radiant energy reflected by a body to that incident upon it.
ϵ, α, ρ	Emissivity, absorptivity, reflectivity	Special cases of emissance, absorptance and reflectance; each is a fundamental property of a material that has an optically smooth surface and is sufficiently thick to be opaque.
τ	Transmittance	The ratio of the radiant energy transmitted through a body to that incident upon it.
τ	Transmissivity	Transmittance for a unit thickness sample.

ϵ, α, ρ, and τ require additional qualifications for precise definition. The terms total and spectral, and the terms hemispherical, normal, and directional, are used and are indicated by subscripts, as illustrated here for emissance, ϵ. In each case the emissance is the ratio of radiation from a surface, as a consequence of its temperature, to that from a blackbody at the same temperature. The subscript indicates the way(s) in which this radiation is limited as to wavelength and/or direction.

ϵ_λ	Spectral emissance	Ratio of spectral radiancy (or monochromatic radiancy at a given wavelength) from a body to that of a blackbody.
ϵ_T	Total emissance	Ratio of total radiancy from a body to that of a blackbody.
ϵ_H	Hemispherical emissance	Ratio of radiancy from a body to that of a blackbody.
ϵ_θ	Directional emissance	Ratio of steradiancy from a body to that of a blackbody.
ϵ_N	Normal emissance	The special case of directional emissance when the emissance is in a direction normal to the surface.

Therefore, precise definitions of emissance will take the following nomenclature:

ϵ_{TH}	Total hemispherical emissance
ϵ_{TN}	Total normal emissance
$\epsilon_{\lambda H}$	Spectral hemispherical emissance
$\epsilon_{\lambda N}$	Spectral normal emissance

*These definitions follow the recent recommendation of NBS personnel who propose the suffix *-ance* for specimen properties and *-ivity* for intrinsic material properties. Emissance follows the suggestion of Judd to avoid confusion with a power flux. The more recent OSA recommendation would change emissance to emittance.

2.2. Kirchhoff's Law: Emissivity and Blackbodies

Kirchhoff's radiation law is

$$\frac{W}{\alpha} = W_{bb} = W^0 \tag{2-4}$$

Kirchhoff's law can also be stated as:

$$\epsilon = \alpha \tag{2-5}$$

Thereby ϵ is W/W_{bb}.

For the limiting conditions of an opaque material ($\tau = 0$) to which Kirchhoff's law applies:

$$\epsilon = 1 - \rho \tag{2-6}$$

A blackbody is a perfect absorber and therefore can be characterized by:

$$\alpha = \epsilon = 1, \quad \rho = 0, \quad \tau = 0 \tag{2-7}$$

2.3. Radiation Laws

Planck's law is the basis for almost all radiometric considerations; other expressions are derivable from it.

2.3.1. Planck's Law. The blackbody spectral energy density for unpolarized radiation is given as

$$W_\lambda = 2\pi c^2 h \lambda^{-5} (e^{hc/\lambda kT} - 1)^{-1}$$

$$W_\lambda = \frac{cc_1}{4} \lambda^{-5} (e^{c_2/\lambda T} - 1)^{-1} \tag{2-8}$$

This pair of equations also defines c_1 and c_2 as

$$c_1 = 8\pi ch$$
$$c_2 = ch/k \tag{2-9}$$

The literature sometimes contains alternative definitions for c_1 that take geometrical factors into account. Usually $c_1 = 8\pi ch$ for the energy density expression, $= 2\pi c^2 h$ for radiant emittance in a hemisphere, and $= 2c^2 h$ for radiance. The first ($c_1 = 8\pi ch$) will be always used here.

W_λ can be related to u_λ as follows:

$$u_\lambda = 4W_\lambda/c \tag{2-10a}$$

where W_λ is the radiant emittance in a uniform enclosure;

$$u_\lambda = 2W_\lambda/c \tag{2-10b}$$

where W_λ is the radiant emittance into a Lambertian hemisphere.

Wavelength and frequency are related as follows:

$$u_\nu = u_\lambda \frac{d\lambda}{d\nu} = \frac{-\lambda^2}{c} u_\lambda = \frac{-c}{\nu^2} u_\lambda \tag{2-11}$$

Another useful form for W_λ is

$$W_\lambda d\lambda = W_\nu d\nu = \frac{2\pi(kT)^4}{c^2h^3}\left(\frac{x^3dx}{e^x - 1}\right) \tag{2-12}$$

$$x = \frac{h\nu}{kT} \tag{2-13}$$

Also,

$$\pi N_\lambda = W_\lambda \tag{2-14}$$

$$4\pi N_\lambda = cu_\lambda \tag{2-15}$$

2.3.2. Quantum Rates in Blackbody Radiation. The energy of a quantum is

$$U = h\nu = \frac{hc}{\lambda} = \frac{(20 \times 10^{-20} \text{ joule } \mu)}{\lambda} = \frac{(1.2398 \text{ ev } \mu)}{\lambda} \tag{2-16}$$

The monochromatic flux n is

$$n_\nu = W_\nu/h\nu = 2\pi c^{-2}\nu^2(e^{h\nu/kT} - 1)^{-1} \text{ photon cm}^{-2} \text{ cps}^{-1} \tag{2-17}$$

$$n_\lambda = W_\lambda\lambda/hc = 2\pi c\lambda^{-4}(e^{hc/\lambda kT} - 1)^{-1} \text{ photon cm}^{-3} \tag{2-18}$$

The mean square fluctuation rate is

$$\overline{(\Delta n)^2} = n\left(1 + \frac{1}{e^{h\nu/kT} - 1}\right) \tag{2-19}$$

When $e^{h\nu/kT}$ is large enough, the classical result is obtained:

$$(\Delta n)^2 = n \tag{2-20}$$

2.3.3. Stefan-Boltzmann Law. When the Planck equation is integrated over all wavelengths the Stefan-Boltzmann expression is obtained:

$$W = \sigma T^4 \tag{2-21}$$

where

W is total radiant emittance, w cm^{-2}

σ is the Stefan-Boltzmann constant, 5.67×10^{-12} w cm^{-2} (°K)$^{-4}$

T is absolute temperature

2.3.4. Rayleigh-Jeans and Wien Laws. These two expressions antedated Planck's formulation. They are incorrect but sometimes give useful approximations.

The Rayleigh-Jeans equation is

$$u_\lambda = c_1\lambda^{-5}(\lambda T/c_2) = 8\pi kT\lambda^{-4} \tag{2-22}$$

The Wien expression is

$$u_\lambda = c_1\lambda^{-5}\, e^{-hc/\lambda kT} = 8\pi c\lambda^{-5}\, e^{-hc/\lambda kT} \tag{2-23}$$

2.3.5. The Wien Displacement Law. This simple expression tells where the peak of the radiation curve falls at any given temperature:

$$\lambda_{max}T = 2897.9\mu \text{ (°K)} \tag{2-24}$$

2.3.6. Maximum Difference Expression. An expression for the wavelength at which the maximum monochromatic radiation difference for a given temperature difference occurs is found by writing

$$\frac{\partial^3 W}{\partial^2\lambda\, \partial T} = 0 \tag{2-25}$$

The resultant condition is

$$\lambda_{max\ diff.} T = 2404\mu\ (°\mathrm{K}) \tag{2-26}$$

2.4. Blackbody Slide Rules

Rules have been devised for rapid, fairly accurate calculations of radiometric quantities.

2.4.1. The General Electric Rule [6]. This rule, designated GEN-15C, is available from the General Electric Company, 1 River Road, Schenectady, New York; it costs about one dollar. Calculations which can be made on the rule are as follows (see Fig. 2-1):

1. Conversions of temperatures among Celsius, Kelvin, Fahrenheit, and Rankine by setting the temperature on one scale and reading it on another – scales *ABKL*.

2. Multiplication by the use of standard *C* and *D* log scales.

3. Total blackbody radiant emittance by setting the temperature of the blackbody source on a temperature scale and reading on the *E* scale (w cm^{-2}). An emissivity scale associated with the *E* scale permits direct calculation for graybodies; read the value on the *E* scale under the appropriate emissivity.

4. Incremental blackbody radiant emittance $W_{\Delta\lambda}$ at maximum. The power density for a 1-μ bandpass can be read directly from the $W_{\lambda max}$ or *F* scale.

5. The ratio of W_λ at any wavelength λ to that at λ_{max}, $W_\lambda/W_{\lambda max}$. The temperature is set on a temperature scale; then $W_\lambda/W_{\lambda max}$ is read from the *G* scale opposite the desired λ on the *H* scale. Thus one can find $W_{\lambda max}$ for a given λ on the $W_{\lambda max}$ scale and then calculate the value of W_λ at any wavelength on the $W_\lambda/W_{\lambda max}$ scale.

6. The blackbody radiation in any spectral interval. Set the temperature scale at the appropriate temperature. Then on the $W_{0-\lambda}/W_{0-\infty}$ or *J* scale read the percentage radiation that lies below a particular wave length λ_1 (on the *I* scale). Do the same for λ_2, and subtract.

7. Conversion of range in nautical miles to range in centimeters with the aid of a straight edge, and vacuum calculation of irradiance. These can be made with the *QRST* scales.

8. Conversion from w $in.^{-2}$ to Btu ft^{-2} hr^{-1}.

9. Number of photons sec^{-1} cm^{-2} from a blackbody at index temperature.

Useful constants and other combinations of these calculations are also available.

2.4.2. The Block Rule. This rule, available from Block Associates Inc., 385 Putnam Avenue, Cambridge 39, Massachusetts, at cost of about two dollars, consists of two curves of blackbody spectral radiance (Fig. 2-2); one is plotted as a function of wavelength and the other as a function of wavenumber. By shifting the curves along the lines marked, blackbody curves for any temperature can be obtained. The curves are made to cross the index line at the appropriate temperature. Then the values for the spectral radiance (unfortunately designated as I_λ and I_ν) are read from the right-hand scales. On the reverse side a curve of $\Delta\nu$ and $\Delta\lambda$ *vs* λ and ν is given. Thus, if one knows a spectral resolution is $\Delta\nu$ at a given ν he can find the resolution $\Delta\nu$ at that ν and the corresponding λ. Of course, calculations from $\Delta\lambda$ to $\Delta\nu$ can also be

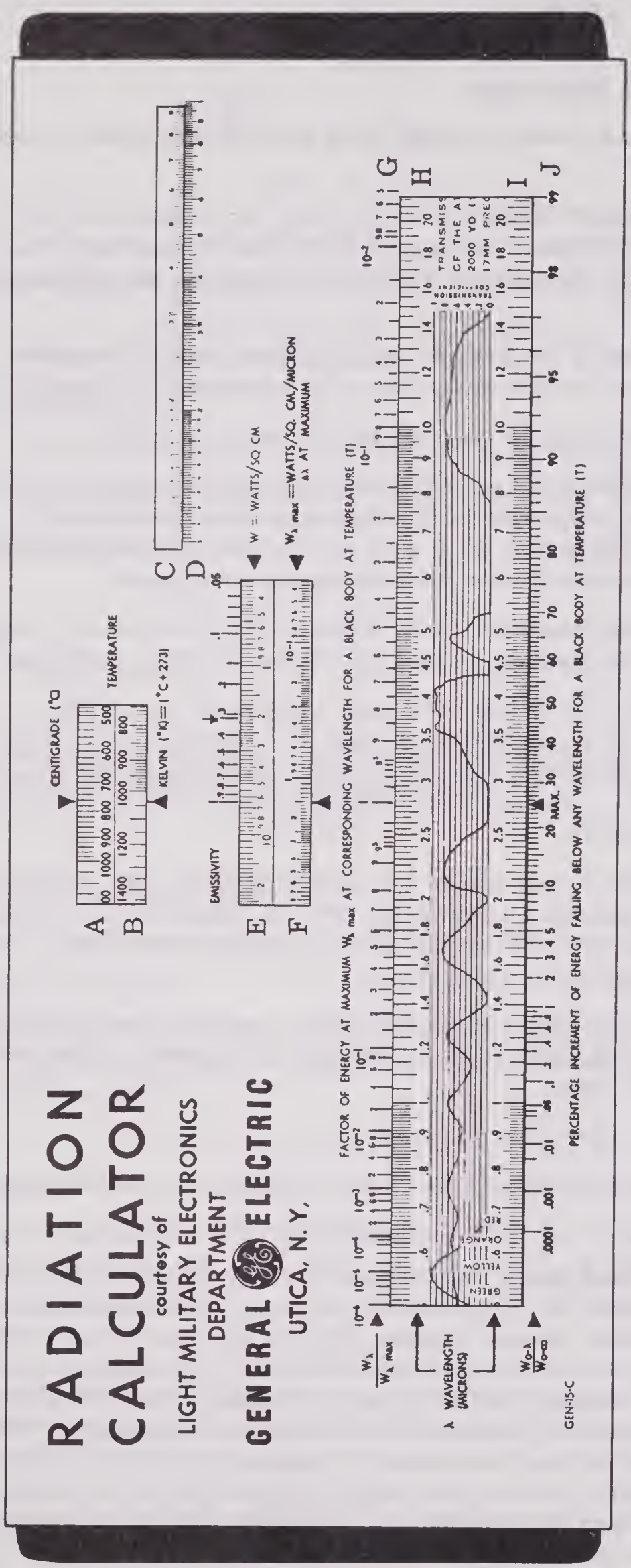


FIG. 2-1. The GE slide rule. [General Electric Co., Light Military Electronics Dept., Utica, N.Y.]

PLANCK'S EQUATION.

$$W_\lambda = C_1 \lambda^{-5} \left(e^{\frac{C_2}{\lambda T}} - 1 \right)^{-1}$$

WIEN'S DISPLACEMENT LAW:

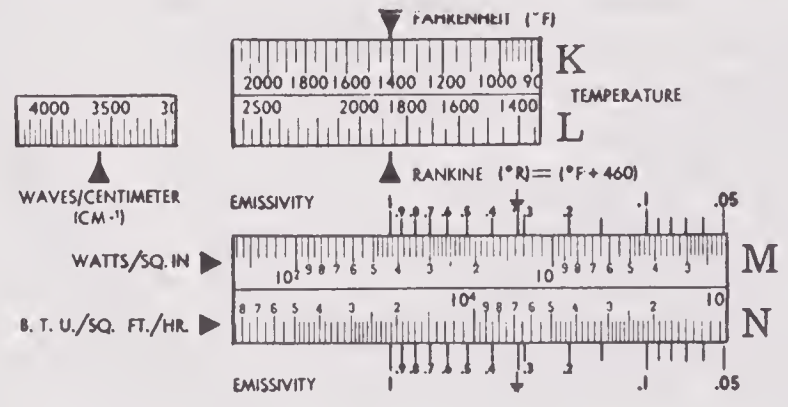
$$\lambda_m = \frac{o}{T}$$

STEFAN-BOLTZMAN LAW

$$W = \epsilon \sigma T^4 \text{ or } W = \epsilon \sigma (T^4 - To^4)$$

STEFAN-BOLTZMANN CONSTANT (σ)

5.6686 x 10^{-5} ergs cm^{-2} deg^{-4} sec^{-1}
5.6686 x 10^{-12} watts cm^{-2} deg^{-4}
1.354 x 10^{-12} cal cm^{-2} deg^{-4} sec^{-1}
3.657 x 10^{-11} watts in^{-2} deg^{-4}
1.797 x 10^{-8} Btu ft^{-2} deg^{-4} hr^{-1}
5.267 x 10^{-12} Kwh ft^{-2} deg^{-4} hr^{-1}
4.529 x 10^{-9} Kcal ft^{-2} deg^{-4} hr^{-1}



SYMBOLS AND PHYSICAL CONSTANTS

W_λ = RADIANT FLUX PER UNIT AREA PER UNIT INCREMENT OF WAVELENGTH = watts/cm^2/cm $\Delta\lambda$ or watts/cm^2/micron $\Delta\lambda$
W = TOTAL RADIANT FLUX EMITTED PER UNIT AREA
T = ABSOLUTE TEMPERATURE OF RADIATING BODY (°K)
T_o = ABSOLUTE TEMPERATURE OF SURROUNDINGS (°K)
λ = WAVELENGTH IN CENTIMETERS OR MICRONS
λ_m = WAVELENGTH IN MICRONS OF MAXIMA OF BLACK BODY CURVE
o = CONSTANT FOR BLACK BODY = 2897.9 MICRON DEGREES
$C_1 = 2\pi c^2 h = 3{,}7413 \times 10^{-12}$ watt cm^2
$C_2 = hc/k = 1.4388$ cm degree
$e = 2.71828$ NAPERIAN BASE
ϵ = EMISSIVITY FACTOR (BLACK BODY = 1)
σ = STEFAN-BOLTZMANN CONSTANT (DEG = K)
c = VELOCITY OF LIGHT = 2.99793×10^{10} cm/sec = 2.99793×10^{14} micron/sec
h = PLANCK'S CONSTANT = 6.6252×10^{-34} watt sec^2
k = BOLTZMANN'S CONSTANT = 1.38042×10^{-23} watt sec/degree
$1/\pi$ = TOTAL RADIANT FLUX PER UNIT SOLID ANGLE (CENTRAL STERADIAN)

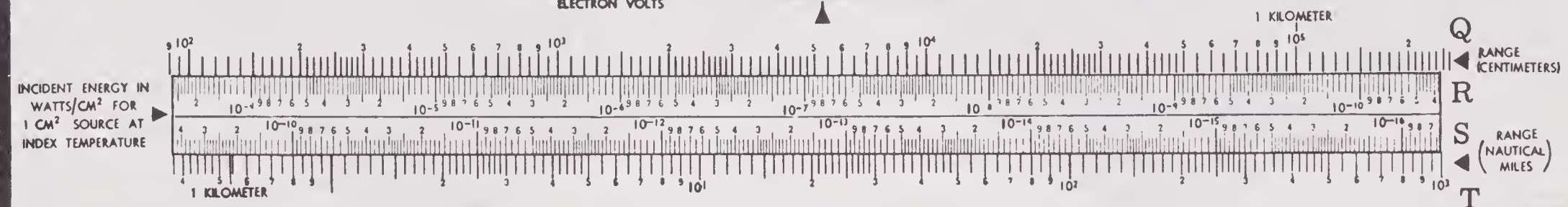


FIG. 2-1 *(Continued).* The GE slide rule. [General Electric Co., Light Military Electronics Dept., Utica, N.Y.]

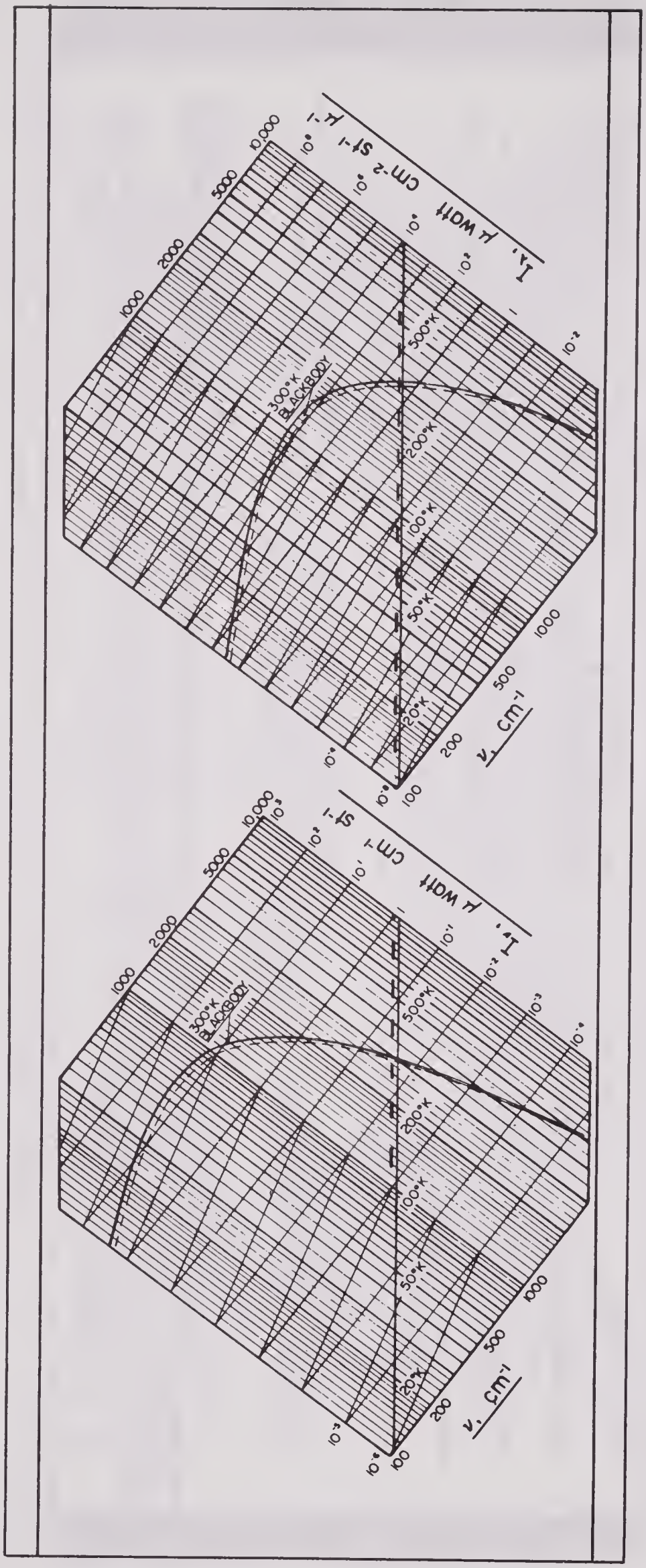


FIG. 2-2. The Block slide rule. [Block Associates, Inc., Cambridge, Mass.]

INSTRUCTIONS

The Radiation Calculator depends upon the fact that curves of Planck's function on log-log coordinates are congruent figures and that a Planck curve for one black-body temperature becomes a Planck curve for another temperature by a parallel translation.

Thus, one Planck curve on the cursor of the calculator can be set to show the irradiance at any one source temperature. Source temperatures are indicated along a line parallel to the cursor edges and diagonal to the log-log grids. Sample Planck curves for the irradiance at 300°K have been drawn in, dashed, on both the grid of I_ν vs. ν and I_λ vs. ν. from about 50 - 500°K. To enhance setting accuracy, note that the temperature scale registers with the wavenumber, or ν, scale. Other temperature ranges are available by multiplying the coordinate scales by factors of ten, for example:

Temperature Range	I_λ	I_ν	ν
500 - 5000°K	x 10^5	x 10^3	x 10

Values of the irradiance, I_ν or I_λ, are only marked each decade. Increased reading accuracy is obtained by projecting the value of irradiance leftwards until it meets one of the diagonal interpolation lines. The location of the irradiance value within its decade is then read off the ν scale.

The units of irradiance are μ watt $cm^{-2}st^{-1}\mu^{-1}$ for I_λ; and μ watt $cm^{-2}st^{-1}cm$ = watt $cm^{-1}st^{-1}$ for I_ν. These two scales are connected by the relation $I_\nu = (\lambda/k)I_\lambda$, where $k \equiv \lambda\nu = 10^4$. Values of k for other units of I_λ are tabulated on the Spectral Resolution Chart.

Block Associates, Inc.

385 PUTNAM AVENUE

CAMBRIDGE 39, MASS.

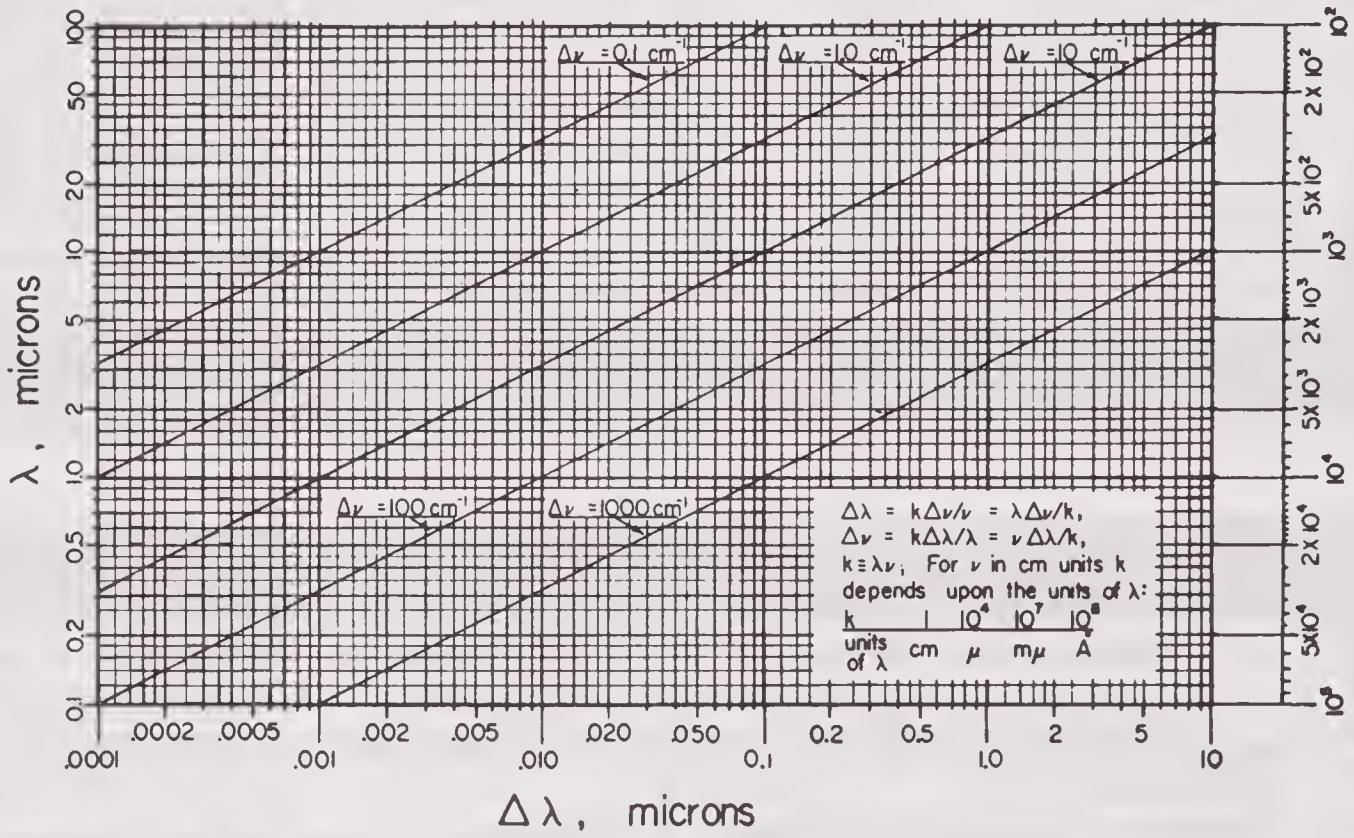


FIG. 2-2 *(Continued).* The Block slide rule. [Block Associates, Inc., Cambridge, Mass.]

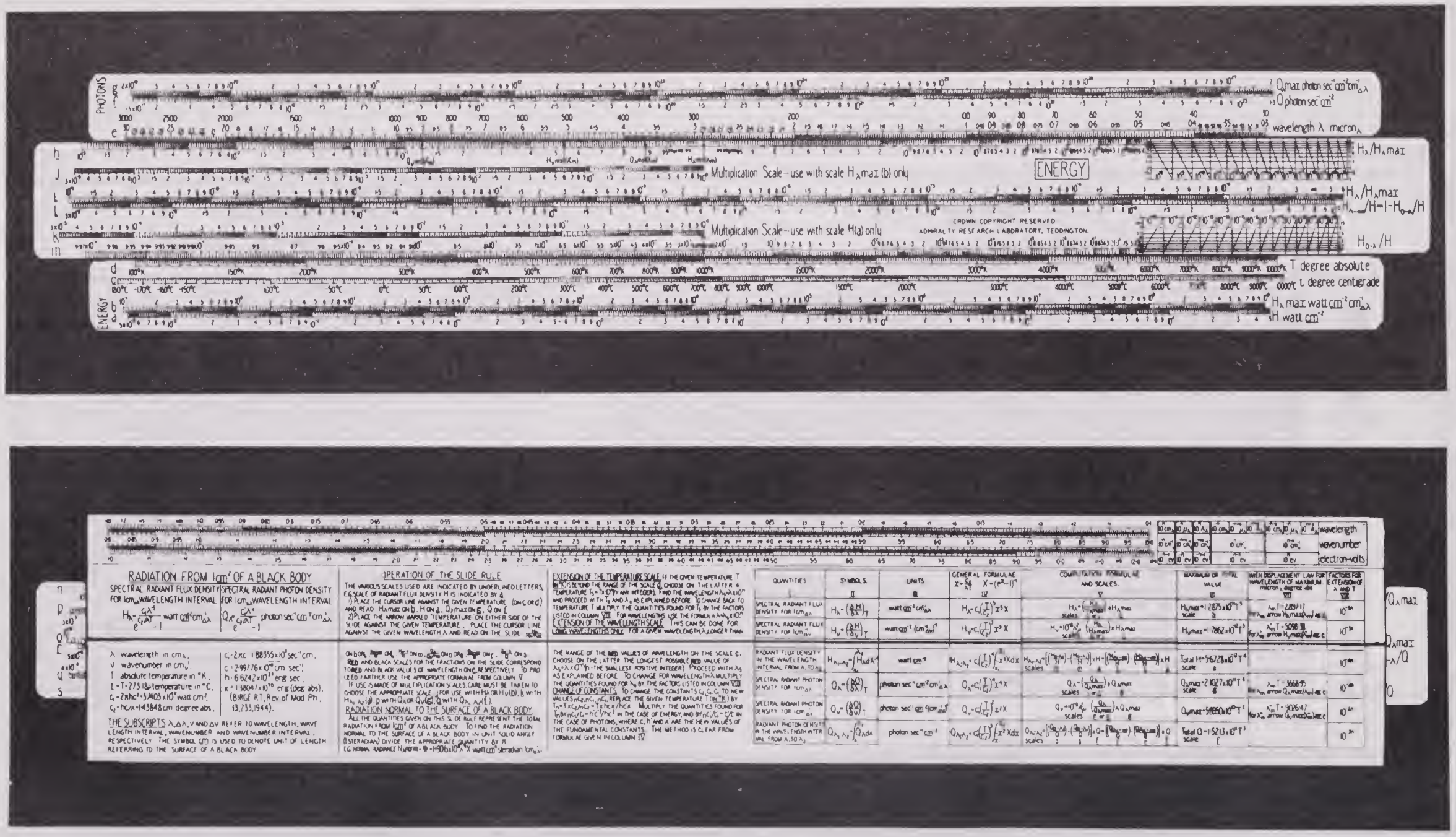


FIG. 2-3. The ARL slide rule. [Admiralty Research Laboratories, Teddington, Middlesex, England.]

made. Although the curves are correct, there are slight errors in the text: the table which is an insert to the spectral resolution calculator should have

$$\Delta\lambda = \frac{-k\,\Delta\nu}{\nu^2} = -\frac{\lambda\,\Delta\nu}{\nu} \tag{2-27}$$

$$\Delta\nu = \frac{-k\,\Delta\lambda}{\lambda^2} = -\frac{\nu\,\Delta\lambda}{\lambda} \tag{2-28}$$

A useful way to remember this relation is

$$\frac{\Delta\lambda}{\lambda} = \frac{-\Delta\nu}{\nu} \tag{2-29}$$

2.4.3. The Admiralty Research Laboratory Rule [7]. This more precise, more expensive (about $150) rule provides essentially the same calculations as the General Electric rule. It can be obtained from: (1) A. G. Thornton Company, Ltd., P.O. Box 3, Wythenshawe, Manchester, England; (2) Jarrell Ash Company, Boston, Massachusetts; or (3) International Scientific and Precision Instrument Company, Inc., 910 Seventeenth Street N.W., Washington 6, D.C. The symbol H, rather than W, is used for flux density (possibly indicating irradiance). The following calculations can be made (Fig. 2-3):

1. The total irradiance (w cm^{-2}) for a given temperature T can be read from scale a under a hairline set at temperature T on scale c (°C) or d (°K).

2. The spectral irradiance for a 1-cm wavelength spectral bandwidth at the maximum of the curve $H_{\lambda max}$ (w cm^{-2}) can be read from scale b under the hairline for the same setting of T on scales c or d.

3. Similar quantities in terms of the number of photons Q can be obtained on scale f for Q (photons sec^{-1} cm^{-2}) and on scale g for $Q_{\lambda max}$ (photons sec^{-1} cm^{-2} cm^{-1}).

4. The ratio of H_λ to $H_{\lambda max}$ can be found by placing the arrow marked TEMPERATURE at the appropriate temperature and using the hairline to find $H_\lambda/H_{\lambda max}$ on scale a for λ on scale i. Similarly, $H_{\lambda-\infty}/H$ can be read on scale i, $H_{0-\lambda}/H$ on scale m, $Q_\lambda/Q_{\lambda max}$ on scale n, $Q_{\lambda-\infty}/Q$ on scale r, or $Q_{0-\lambda}/Q$ on scale s.

5. Scale e is the wavelength scale on which one can also read wavelengths at which H_λ, Q_λ, H_ν, and Q_ν have their maximum values.

The temperature scale runs from 100°K to 10,000°K, but if higher or lower temperatures are desired, the rule can be extended by the use of the multiplication tables k and j. The instructions given here and more are provided on the reverse side of the rule, which also gives wavelength, wavenumber, and energy scales (but no hairline for conversion!).

2.5. Blackbody Curves [8]

Figures 2-4, 2-5, and 2-6 provide information about the spectral distribution of blackbody radiation. The first of these is a linear curve with wavelength as the abscissa and W_λ as the ordinate. The second is semilogarithmic but for a different range of temperatures. The third is a curve plotted on a logarithmic wavelength scale and linear emittance scale in terms of the variable λT.

Another useful curve is a log-log plot of the Planck equation. The shape of the curve is identical for all T and need only be shifted along the line representing the Wien displacement law. Such curves are shown in Figs. 2-7 and 2-8. The straight lines are the "sliding lines." Every blackbody curve for any temperature can be obtained

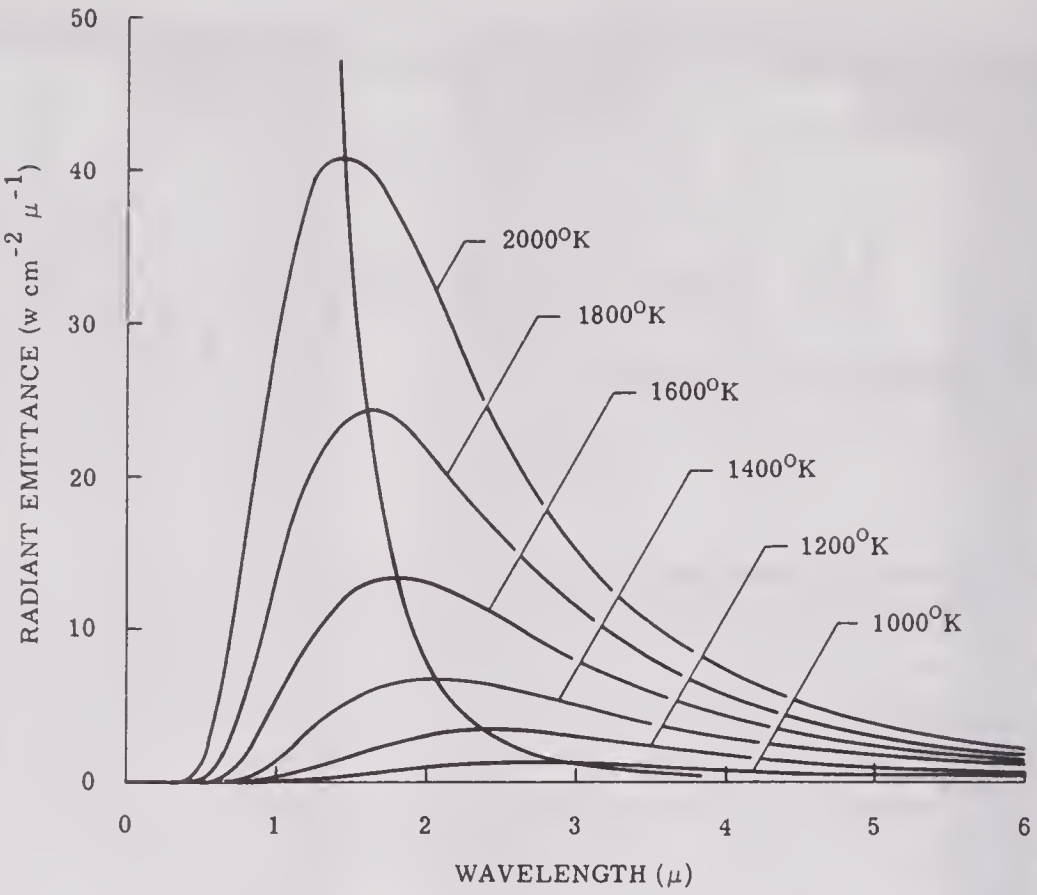


FIG. 2-4. Blackbody curves, 1000°K to 2000°K.

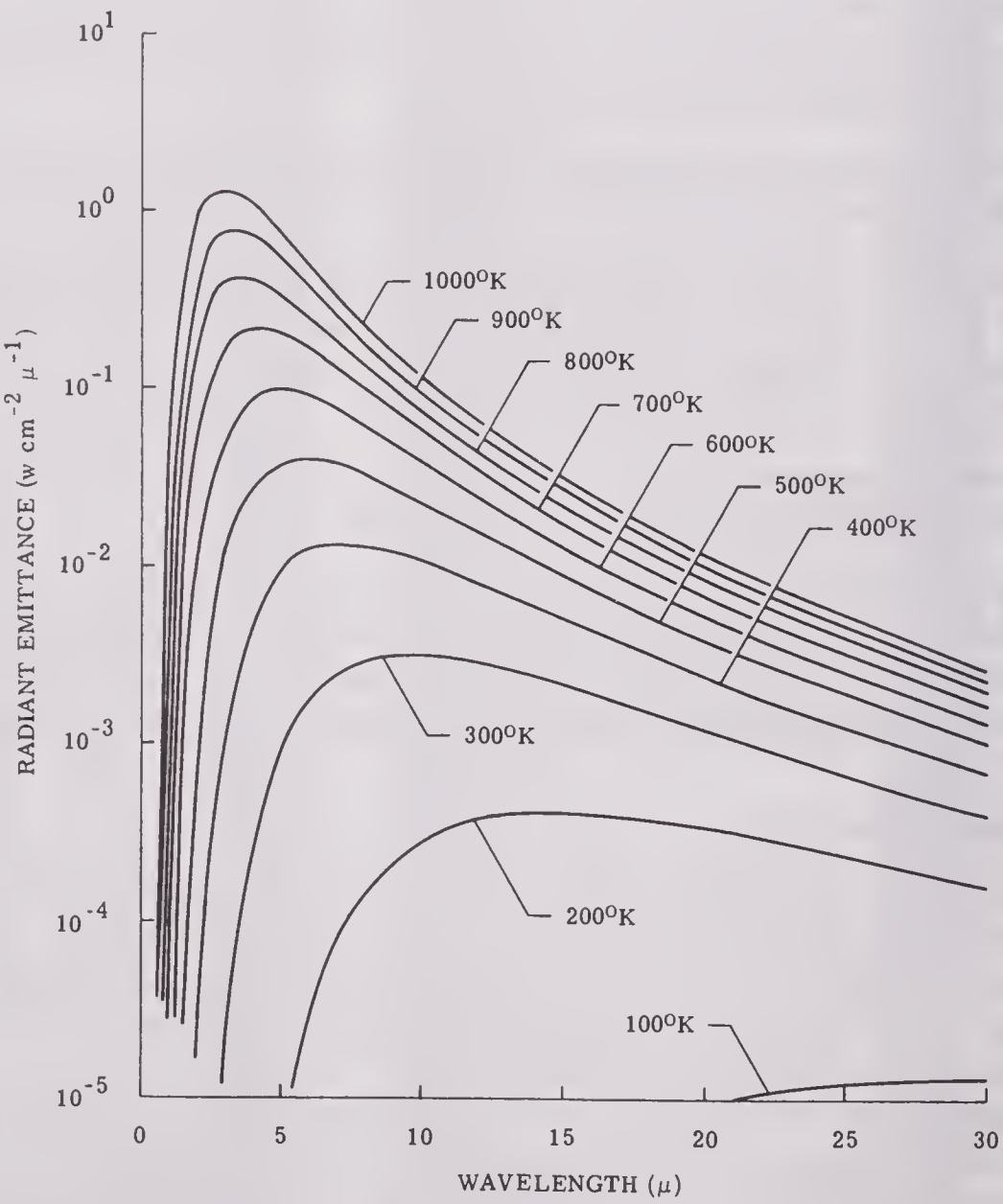


FIG. 2-5. Blackbody curves, 100°K to 1000°K.

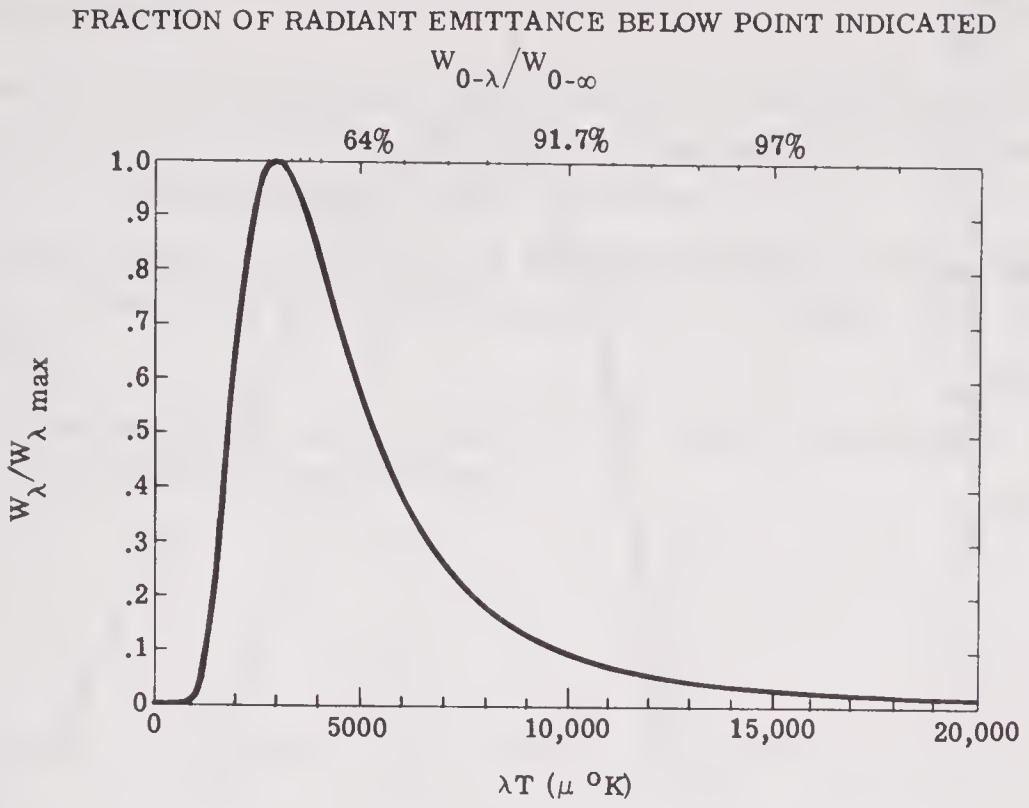


FIG. 2-6. $W_\lambda/W_{\lambda\,max}$ *vs* λT for a blackbody.

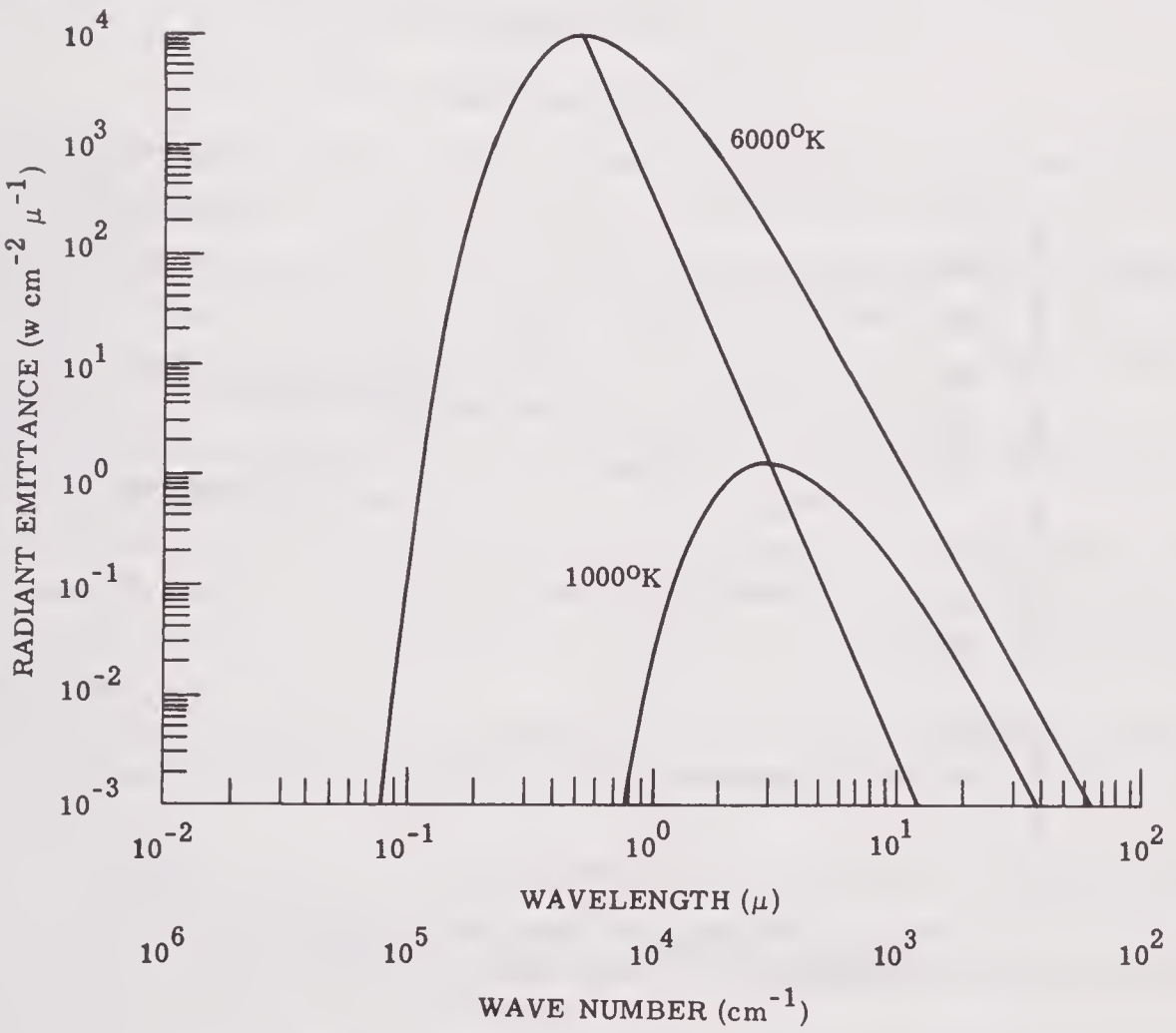


FIG. 2-7. W_λ *vs* λ and ν.

by moving a curve of the same shape along this line. Thus a "do-it-yourself" slide rule can be constructed by putting an overlay on this figure, tracing the curve and the line, and placing an index marker at 6000°K, the temperature of the top curve. Then by keeping the lines overlapped and setting the index marker at the desired temperature, the template becomes the blackbody curve for that temperature. Figure 2-8 is the same sort of curve but for a different temperature region. These are similar to the Block slide rule.

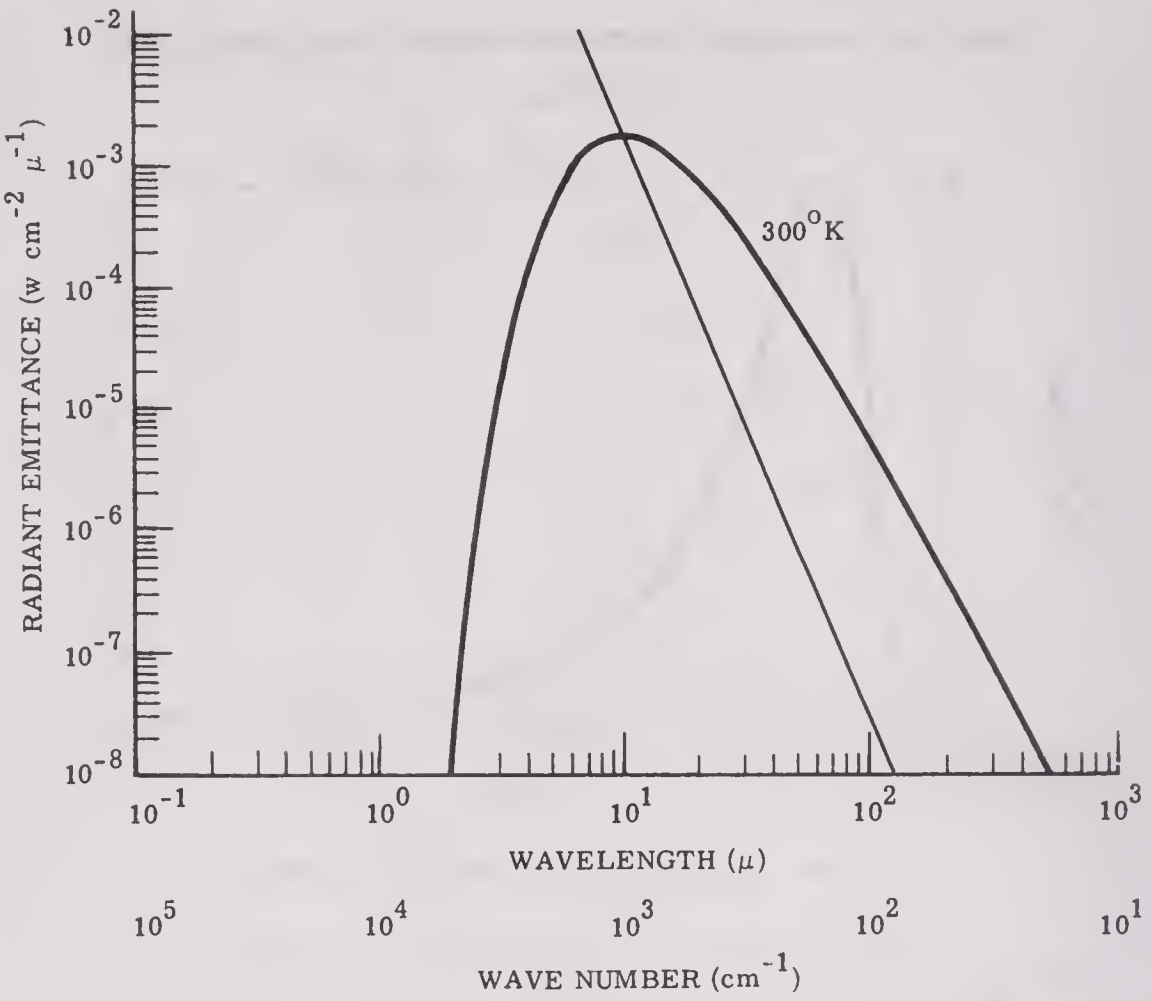


FIG. 2-8. W_λ *vs* λ and ν.

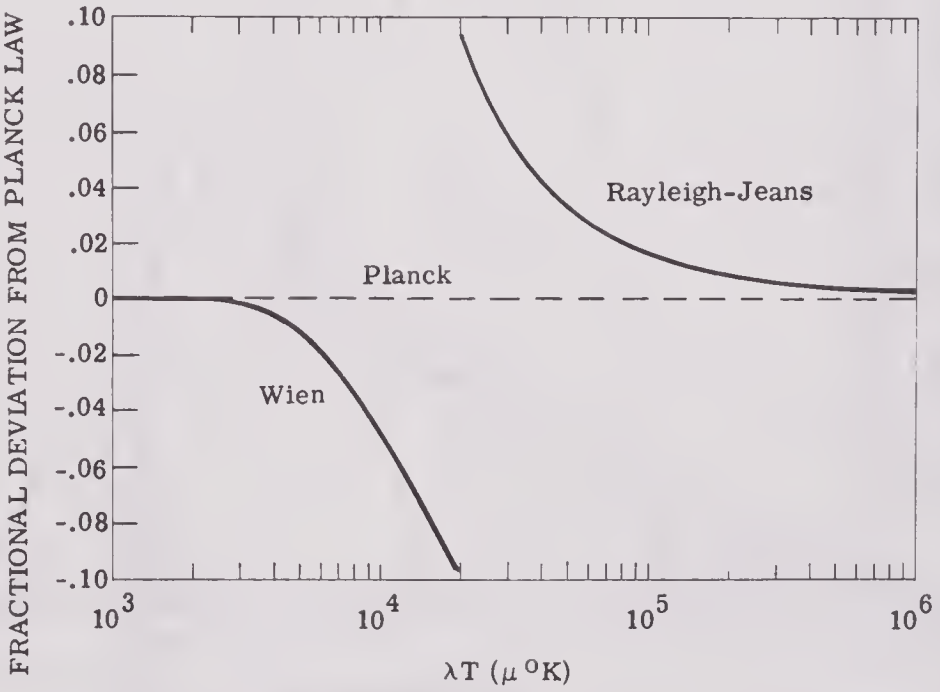


FIG. 2-9. Deviations of Wien and Rayleigh-Jeans expressions from Planck's law.

Often blackbody calculations can be made on the basis of either the Rayleigh-Jeans or the Wien expression. Thus, when $hc/\lambda kT$ is large,

$$N_\lambda = 2c^2h\lambda^{-5}\, e^{-hc/\lambda kT} \tag{2-30}$$

When $hc/\lambda kT$ is small,

$$N_\lambda = 2ck\lambda^{-4}T \tag{2-31}$$

Figure 2-9 provides information about errors inherent in using these expressions.

2.6. Blackbody Tables

For very careful work it is necessary to use tables of blackbody functions with their additional precision and attendant difficulties. Since each set of tables is a thick book in itself, only references to these works are given here.

S. A. Golden, *Spectral and Integrated Blackbody Radiation Functions,* Research Report 60-23, Rocketdyne Division, North American Aviation, Inc., Canoga Park, Calif. (1960) [9].

Table I provides $W_\lambda/W_{\lambda max}$ and $W_{0-\lambda}/W$ and their first derivative as functions of $c_2/\lambda T$. The intervals are as follows:

$c_2/\lambda T$	$\Delta c_2/\lambda T$
0-2	0.01
2-5	0.02
5-10	0.05
10-25	0.10
25-50	0.20

Table II provides $W_{\lambda max}$ and W as functions of T from 0°K to 10,000°K in 10° K intervals. Values of the radiation constants are

$$c_1 = 2\pi hc^2 = 3.7413 \times 10^{-5} \text{ erg cm}^2 \text{ sec}^{-1}$$

$$c_2 = hc/k = 1.4388 \text{ cm °K}$$

M. Pivovonsky and M. Nagel, *Tables of Blackbody Radiation Functions,* Macmillan; New York (1961) [10].

Table I is a tabulation of (1) N_λ *vs* λ and T; (2) the ratio $N_\lambda(\lambda T_1)/N$ (0.560 μ, T_1); (3) $N_{0-\lambda}/N$. These are tabulated to five significant figures for $\lambda = 0.2$ μ to 0.590 μ in 0.005-μ intervals and between 0.590 μ and 1.2 μ in 0.01-μ intervals from 800 to 40,000°K in intervals varied to meet the needs of the range.

Table II continues Table I for $\lambda = 1.1$ μ to 1100 μ at temperatures from 20 to 13,000°K with four-figure accuracy.

Table III includes: (1) $N_\lambda/N_{\lambda max}$; (2) a restatement of the wavelength ratios of Table I and II; and (3) a function for computing derivatives of the Planck function. These are plotted for λT from 0.01 to 0.99 μ °K. Procedures for evaluation at higher values of λT are given.

Table IV has: N, $N_{\lambda max}$, λ_{max} for temperatures from 1000° to 2500°K at 2°K intervals, from 2500 to 5500°K in 5°K intervals and from 5500° to 10,000°K in 10°K intervals.

Table V repeats Table IV but for wavenumbers, and Table VI repeats Table IV for reciprocal temperatures.

Table VII gives luminance from 800° to 1796°K in 4°K intervals and a table of luminance and chromaticity coordinates.

Table VIII is a temperature correction table – for revised physical constants.

M. Czerny and A. Walther, *Tables of the Fractional Function for the Planck Radiation Law,* Springer-Verlag, Berlin (1961) [11].

Deals with $W_{0-\lambda}/W$ plotted *vs* $\lambda T/c_2$. The tables are independent of c_2. It also includes the first and second derivatives of this function and

$$\frac{W_{0-\lambda}}{W} + \frac{1}{4}\frac{\lambda T}{c_2}\frac{d}{d\lambda}\left(\frac{W_{0-\lambda}}{W}\right)$$

and its first derivative.

Other tables in print are as follows:

Parry Moon, *J. Opt. Soc. Am.* **38,** 291 (1948) [12].
A. N. Lowen and G. Blanch, *J. Opt. Soc. Am.* **30,** 70 (1940) [13].
E. Jahnke and F. Emde, *Tables of Functions,* Dover, New York (1945) [14].
A. G. DeBell, Rocketdyne Research Report 59-32, Rocketdyne Division of North American Aviation, Inc., Canoga Park, Calif. (1959) [15].

Still others exist for special uses: C. C. Ferriso of Convair prepared a set for flame calculations (in terms of wavenumber) and S. Twomey of NASA prepared a set (180° to 315°K and 20 cm^{-1} to 3400 cm^{-1}), mostly useful for meteorological work.

2.7. Radiation Geometry

The differential solid angle is (see Fig. 2-10):

$$d\Omega = \frac{r^2 \sin\theta \, d\theta \, d\phi}{r^2} \tag{2-32}$$

The solid angle of a sphere is 4π sr.

The solid angle of a hemisphere is 2π sr.

For small angles the solid angle is the product of the plane angles of two sides of an area, and is equal to the area divided by the distance squared.

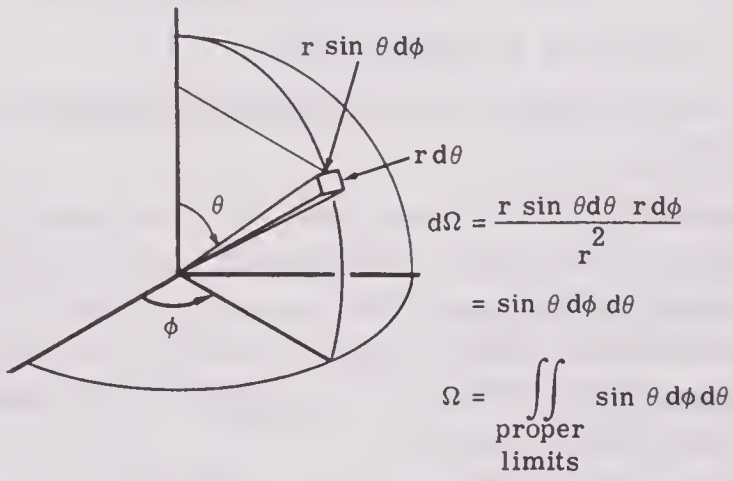


FIG. 2-10. Radiation geometry.

2.7.1. Lambertian Sources. When radiance is defined, as in Table 2-1, as

$$N \equiv \frac{\partial^2 P}{\cos\theta \, \partial A \, \partial\Omega} \qquad \text{w cm}^{-2}\ \text{sr}^{-1} \tag{2-33}$$

a Lambertian surface has a constant radiance, which is the same in all directions.

In general, when the radiance of a surface at a point is expressed as a function of direction, $N = N(\theta, \varphi)$, the radiant emittance at that point is given by

$$W \equiv \frac{\partial P}{\partial A} = \int N \cos\theta \, d\Omega$$

$$= \int\int N(\theta, \varphi) \cos\theta \sin\theta \, d\theta \, d\varphi \qquad \text{w cm}^{-2} \tag{2-34}$$

where the integration covers the entire solid angle containing the radiation beam of interest. If this is a perpendicular inverted circular cone of half-angle α,

$$W = \int_0^{2\pi} \int_0^{\alpha} N(\theta, \varphi) \cos\theta \sin\theta \, d\theta \, d\varphi \qquad \text{w cm}^{-2} \tag{2-35}$$

If the surface is Lambertian ($N \equiv$ constant),

$$W = 2\pi N \int_0^{\alpha} \cos\theta \sin\theta \, d\theta = \pi N \sin^2 \alpha \qquad \text{w cm}^{-2} \tag{2-36}$$

The radiant emittance into a hemisphere ($\alpha = \pi/2$) from a Lambertian surface is then

$$W = \pi N \qquad \text{w cm}^{-2} \tag{2-37}$$

It is *not*

$$W = 2\pi N \qquad \text{w cm}^{-2} \tag{2-38}$$

2.8. Distributed Radiators

If a volume rather than a surface is the source of radiation, then when pressure and temperature along the path are constant,

$$W_\lambda = W_\lambda{}^0(1 - e^{-a_\lambda x}) \tag{2-39}$$

where W_λ = spectral radiant emittance of the radiator

$W_\lambda{}^0$ = spectral radiant emittance of a blackbody

a_λ = spectral absorption coefficient

x = radiating path of the material

The emissivity of such a partially transparent body is

$$\epsilon_\lambda = 1 - e^{-ax} \tag{2-40}$$

2.9. Selective Radiators

The radiation from selective radiators can be calculated only by complicated processes. Below are listed some of the useful relations. Penner [3] is a good source for the theory.

2.10. Directional Reflectance and Emissivity

The basic interrelationships among the six quantities, normal and diffuse reflectance and transmittance, and hemispherical and directional emissivity, are concisely stated in an appendix to a paper by Richmond beginning on page 151 of [16]. These relationships are not easily visualized, and some readers may find the following alternative approach and terminology helpful. Only opaque bodies of zero transmittance are treated.

The radiance N_i, a function of both position and direction, is incident on the surface of an opaque body where some of the radiation is absorbed and the rest is reflected (as used here, *reflected* includes diffuse reflectance or scattering) to form a second radiation field, where the radiance N_r of the reflected radiation is also a function of position and direction. N_r is directly proportional to N_i in the sense that, if the value of N_i is multiplied by a constant that is independent of position and direction, the resulting values of N_r will all be multiplied by the same constant factor. However, it will be seen below that the interdependence of the spatial and directional distributions of N_r and N_i is more complex.

The radiant power incident on a particular element δA of the reflecting surface shown in Fig. 2-11, through an elementary beam of solid angle $\delta\Omega_i$, from a direction (θ_i, φ_i), is given by

$$\begin{aligned} \delta P_i(\theta_i, \varphi_i) &= N_i(\theta_i, \varphi_i) \cos\theta_i \, \delta A \delta\Omega_i \\ &= N_i(\theta_i, \varphi_i)\delta\Omega'_i \, \delta A \qquad \text{w} \end{aligned} \tag{2-41}$$

where $\delta\Omega'_i = \cos\theta_i \, \delta\Omega$

$= \sin\theta_i \cos\theta_i \, d\theta_i \, d\varphi_i.$

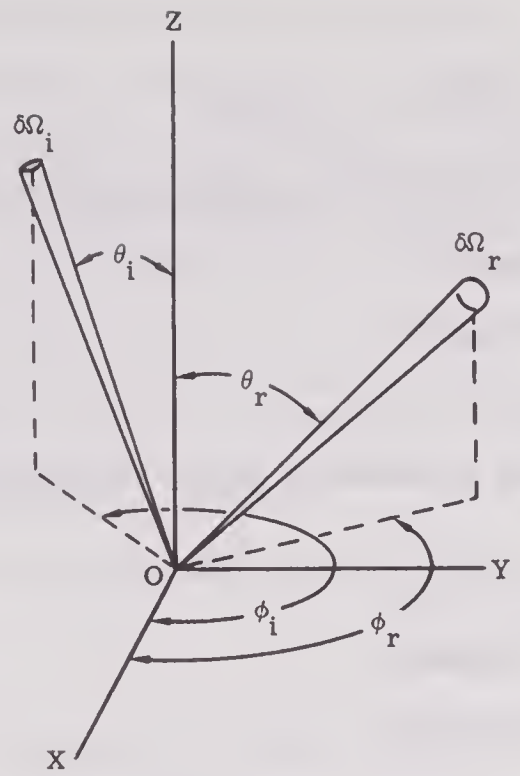


FIG. 2-11. Geometry of incident and reflected elementary beams. The Z axis is chosen along the normal to the surface element at O.

The quantity $\delta\Omega'_i$ is the *projected solid angle* [5,17] of the elementary beam. Correspondingly, the irradiance at δA is

$$\delta H_i(\theta_i, \varphi_i) = N_i(\theta_i, \varphi_i)\delta\Omega'_i \qquad \text{w cm}^{-2} \tag{2-42}$$

Then the radiant intensity of the surface element δA, due to reflection (scattering) of radiation from this incident elementary beam, in the direction (θ_r, φ_r) is

$$\delta J_r(\theta_r, \varphi_r) = \rho'(\theta_i, \varphi_i, \theta_r, \varphi_r) \cos\theta_r \delta P_i(\theta_i, \varphi_i) \qquad \text{w sr}^{-1} \tag{2-43}$$

and the reflected (scattered) radiance

$$\delta N_r(\theta_r, \varphi_r) = \rho'(\theta_i, \varphi_i, \theta_r, \varphi_r)\delta H_i(\theta_i, \varphi_i) \qquad \text{w cm}^{-2}\ \text{sr}^{-1} \tag{2-44}$$

where

$$\rho'(\theta_i, \varphi_i, \theta_r, \varphi_r) \equiv \frac{\delta N_r(\theta_r, \phi_r)}{\delta H_i(\theta_i, \phi_i)} = \frac{\delta N_r(\theta_r, \phi_r)}{N_i(\theta_i, \phi_i)\delta\Omega'_i} \qquad \text{sr}^{-1} \tag{2-45}$$

ρ' is the partial reflectance, or *reflection-distribution function* [18] of the surface element δA for radiation incident from the direction (θ_i, ϕ_i) and reflected (scattered) in the direction (θ_r, ϕ_r). Furthermore, by a reciprocity theorem of wide generality [19, 20] first enunciated by Helmholtz:*

$$\rho'(\theta_1, \phi_1, \theta_2, \phi_2) = \rho'(\theta_2, \phi_2, \theta_1, \phi_1) \qquad \text{sr}^{-1} \tag{2-46}$$

Thus $\rho'(\theta_1, \phi_1, \theta_2, \phi_2)$ is the partial reflectance between the two directions (θ_1, ϕ_1) and (θ_2, ϕ_2), where either direction may be that of the incident elementary beam and the other that of the reflected (scattered) elementary beam.

Hence the radiance at a point of the reflecting surface (taken as the origin for spherical coordinates) in the direction (θ_r, φ_r) and due to reflection (scattering) of all beams of incident radiation is

---

*A search for a proof (in English) of this important theorem also turned up a number of authors who referred to or made use of the theorem in various ways without giving a proof [21-24], including von Helmholtz himself [25], although Planck [26] states, without specific citation, that von Helmholtz "proved" the theorem. DeHoop [20] not only gives a proof (essentially the same as that of Kerr [19]) but also includes an explicit statement of the requisite conditions.

$$N_r(\theta_r, \varphi_r) = \int_0^{2\pi} \int_0^{\pi/2} \rho' N_i \sin \theta_i \cos \theta_i \, d\theta_i \, d\phi_i$$

$$= \int_h \rho' N_i \, d\Omega'_i \qquad \text{w cm}^{-2} \text{ sr}^{-1} \tag{2-47}$$

The following notation is used to designate integration over a hemisphere:

$$\int_h f(\theta, \varphi) \, d\Omega \equiv \int_0^{2\pi} \int_0^{\pi/2} f(\theta, \varphi) \sin \theta \, d\theta \, d\varphi$$

and

$$\int_h f(\theta, \varphi) \, d\Omega' \equiv \int_0^{2\pi} \int_0^{\pi/2} f(\theta, \varphi) \sin \theta \cos \theta \, d\theta \, d\varphi$$

Relation (2-47) is for a particular point, or for the surface element δA at that point. For a more general expression, one must also establish the reflected radiance from other points. When ρ' and N_i are expressed as functions of spatial location (as well as direction) for all points on the reflecting surface, Eq. (2-47) gives the reflected radiance N_r as a function of position for these same points on the reflecting surface, as well as for direction (θ_r, φ_r) at each such point. It is important here to recognize that Eq. (2-47) is written above in coordinates which, for convenience, are specially oriented with respect to the surface element δA. Appropriate adjustments must be made when dealing with irregular surfaces where the direction of the normal changes in going from one surface element to another.

Whether surface irregularities are treated as microscopic [in the sense that their effects are integrated or averaged in the distribution function or partial reflectance $\rho'(\theta_i, \varphi_1, \theta_2, \varphi_2)$], or as macroscopic (in the sense that they may be analyzed into smaller surface elements δA for treatment as above) can be arbitrary, depending on the degree of resolution desired, or can be dependent on circumstances limiting achievable resolution. For example, in examining the reflectance of a highly irregular surface containing deep cavities, such as a piece of volcanic scoria, or a coarse, blackened cellulose sponge in the laboratory, it may be possible to consider the reflectance of different portions of the walls of single cavities (which are then regarded as macroscopic irregularities). But when one studies the possible effects of similar surfaces which may exist on the moon, where such fine detail cannot possibly be resolved by the best telescopes on earth, these are necessarily treated as microscopic irregularities [27,28]. Still more complicated considerations are introduced when microscopic irregularities are small enough to have dimensions of about, or less than, the wavelength of the incident light or other electromagnetic radiation [29-32].

The total reflectance ρ of a surface element δA is defined in general as

$$\rho \equiv \delta P_r / \delta P_i \qquad \text{dimensionless} \tag{2-48}$$

where δP_i is the total radiant power incident (from all directions) on δA, and δP_r is the total resulting reflected radiant power (in all directions). As stated above, the value of ρ depends on the geometry and spectrum of the incident beam of radiation, which may be different in each particular case. Here, for the moment, the primary concern is the geometry. Hence spectral considerations will be eliminated for the remainder of this section by restricting the spectrum of the incident radiation, except where otherwise stated, to a region over which ρ does not change significantly with wavelength. It is then useful to consider some special cases of incident-beam geometry.

If the incident radiation is well collimated, within a small element of solid angle $\delta\Omega_i = \sin\theta_i\, d\theta_i\, d\varphi_i$ from the direction (θ_i, φ_i), the total radiant power incident on δA is

$$\delta P_i = \delta H_i(\theta_i, \varphi_i)\, \delta A \qquad \text{w} \tag{2-49}$$

Then, from Eq. (2-45),

$$\begin{aligned} \delta N_r(\theta_r, \varphi_r) &= \rho'\, \delta H_i(\theta_i, \phi_i) \\ &= \rho'\, \delta P_i/\delta A \qquad \text{w cm}^{-2}\ \text{sr}^{-1} \end{aligned} \tag{2-50}$$

But

$$\begin{aligned} \delta P_r &= \delta A \int_h \delta N_r(\theta_r, \varphi_r)\, d\Omega'_r \\ &= \delta P_i \int_h \rho'(\theta_i, \varphi_i, \theta_r, \varphi_r)\, d\Omega'_r \\ &= \delta P_i \rho_{di}(\theta_i, \varphi_i) \qquad \text{w} \end{aligned} \tag{2-51}$$

where $\rho_{di}(\theta_i, \varphi_i)$ is the (total) directional reflectance for a well-collimated incident beam:

$$\rho_{di}(\theta_i, \varphi_i) \equiv \int_h \rho'(\theta_i, \varphi_i, \theta_r, \varphi_r)\, d\Omega'_r \qquad \text{dimensionless} \tag{2-52}$$

For isotropic surfaces, there is no dependence on the azimuth φ, and Eq. (2-52) simplifies to the frequently recognized dependence on θ: $\rho_{di}(\theta_i, \varphi_i) = \rho_{di}(\theta_i)$. If the well-collimated beam is incident perpendicularly on a plane surface, this becomes the commonly reported normal reflectance $\rho_n = \rho_{di}(0)$. If a point on the surface of an opaque solid is uniformly irradiated from all external directions, *i.e.*, if N_i is a constant, the reflected radiance in the direction (θ_r, φ_r), from Eq. (2-47) is given by

$$\begin{aligned} N_r(\theta_r, \varphi_r) &= N_i \int_h \rho'\, d\Omega'_i \\ &= N_i \rho_{dr}(\theta_r, \varphi_r) \qquad \text{w cm}^{-2}\ \text{sr}^{-1} \end{aligned} \tag{2-53}$$

where

$$\rho_{dr}(\theta_r, \varphi_r) \equiv \int_h \rho'(\theta_i, \varphi_i, \theta_r, \varphi_r)\, d\Omega'_i \qquad \text{dimensionless} \tag{2-54}$$

But, from the reciprocity relation, Eq. (2-46), and Eqs. (2-52) and (2-54),

$$\rho_{di}(\theta_1, \varphi_1) = \rho_{dr}(\theta_1, \varphi_1) = \rho_d(\theta_1, \varphi_1) \qquad \text{dimensionless} \tag{2-55}$$

Thus the (total) directional reflectance $\rho_d(\theta_1, \varphi_1)$ for a well-collimated beam incident from the direction (θ_1, φ_1) is also the ratio between the reflected radiance $N_r(\theta_1, \varphi_1)$ in that same direction and the incident radiance N_i when the surface is uniformly irradiated from all directions (hemispherical irradiation). This relation [Eqs. (2-53) and (2-55)] is the basis for a reflectometry technique described by McNicholas [22].

More important, Eqs. (2-51), (2-53), and (2-55) are the basis for evaluating and equating the directional absorptance and directional emissivity of the surface element δA in a simple relation which has the same form as the Kirchhoff's law relation, Eq. (2-6). If, in Eq. (2-53), the uniform incident radiance N_i is equal to $N_b(T)$, the blackbody radiance (either total or spectral, *i.e.*, in a small wavelength interval at a given wavelength) in an isothermal enclosure at T°K, and if, in fact, the reflecting surface forms the wall of such an enclosure so that it too is at this same temperature, then the radiance in the direction (θ_1, φ_1) from the element of wall surface δA is made up of an emitted radiance and a reflected radiance, as follows:

$$N_e + N_r = \epsilon_d(\theta_1, \varphi_1) N_b(T) + \rho_{dr}(\theta_1, \varphi_1) N_b(T) = N_b(T) \qquad \text{w cm}^{-2}\ \text{sr}^{-1} \tag{2-56a}$$

Similarly, the radiance from the direction (θ_1, φ_1) incident on the element δA is made up of an absorbed incident radiance and a reflected incident radiance (scattered in all directions):

$$N_{ia} + N_{ir} = \alpha_d(\theta_1, \varphi_1)N_b(T) + \rho_{di}(\theta_1, \varphi_1)N_b(T)$$

$$= N_b(T) \qquad \text{w cm}^{-2}\ \text{sr}^{-1} \tag{2-56b}$$

Here, $\epsilon_d(\theta_1, \varphi_1)$ is the directional emissivity (at temperature T) of the element δA for radiation emitted in the direction (θ_1, φ_1) and $\alpha_d(\theta_1, \varphi_1)$ is the absorptance (at T) for radiation incident *from* that direction. Consequently, from Eq. (2-55),

$$\epsilon_d(\theta_1, \phi_1) = 1 - \rho_{dr}(\theta_1, \phi_1)$$

$$= 1 - \rho_{di}(\theta_1, \phi_1) = \alpha_d(\theta_1, \phi_1) \qquad \text{dimensionless} \tag{2-57}$$

Note that equilibrium maintenance with conservation of energy (Kirchhoff's law) by itself would justify only each line of Eq. (2-57) independently, and the Helmholtz Reciprocity law (which is the basis for Eq. (2-46) and, in turn, Eq. (2-55)) must also be invoked in order to equate them to each other and so to relate emissivity for radiation emitted into a given direction to the absorptance for radiation incident *from* that same direction.

In the more familiar form of Kirchhoff's law,

$$\epsilon = 1 - \rho = \alpha \qquad \text{dimensionless} \tag{2-6}$$

directional quantities are not considered. Instead, the total emissivity for radiation emitted in all directions (into a hemisphere) is related to the total reflectance (in all directions into a hemisphere) for uniform incident radiance (from all directions, *i.e.*, from a hemisphere) and to the total absorptance for uniform incident radiance (from all directions, *i.e.*, from a hemisphere). The total reflectance ρ in Eq. (2-6), for uniform incident radiance (N_i = a constant independent of direction) is then

$$\rho = \delta P_r/\delta P_i = \frac{\delta A \int_h N_r\, \delta\Omega'}{\delta A \int_h N_i\, \delta\Omega'} = \frac{N_i \int_h \rho_d(\theta, \varphi)\, d\Omega'}{N_i \int_h d\Omega'}$$

$$= \frac{1}{\pi} \int_h \rho_d(\theta, \varphi)\, d\Omega' \qquad \text{dimensionless} \tag{2-59}$$

The quantities in Eq. (2-6) are those involved in heat-transfer computations where the interest is in the net flow of energy across a bounding surface, involving radiation received, emitted, or reflected in all directions.

Equations (2-6) and (2-57) apply in all cases to spectral radiation (*i.e.*, the radiation in a very small wavelength interval about a specified wavelength) and hence also to any spectral interval in which ρ or ρ_d (and therefore also ϵ or ϵ_d and α or α_d) do not change significantly with wavelength. When thermal equilibrium exists, *i.e.*, when

$$N_i = N_b(T) = \int_0^\infty N_{\lambda b}(T, \lambda)\, d\lambda$$

where $N_{\lambda b}(T, \lambda)$ is the spectral radiance of a blackbody at T°K, they also apply to total radiation (all wavelengths), even though the spectral reflectance varies with wavelength. However, if the spectral reflectance is not a constant *and* the spectral distribution of the incident radiation is arbitrary (nonequilibrium condition), Eq. (2-6)

and (2-57) do not necessarily hold for the total (all wavelengths) reflectance, absorptance, and emissivity.

The significance of the partial reflectance ρ' and directional reflectance ρ_d can be clarified by relating them to the more familiar ideas of diffuse and specular reflectance. First, a perfectly diffuse reflector is one for which $\rho' =$ a constant, so that

$$\rho_d = \rho' \int_h d\Omega' = \pi\rho' \qquad \text{dimensionless} \tag{2-60}$$

Hence ρ_d is also a constant, *i.e.*, the same in all directions, so the total reflectance $\rho = \rho_d = \pi\rho'$ for any arbitrary configuration of incident radiation. Second, a perfectly specular reflector is characterized by

$$N_r(\theta, \varphi \pm \pi) = \rho_d(\theta, \varphi) N_i(\theta, \varphi) \qquad \text{w cm}^{-2}\ \text{sr}^{-1} \tag{2-61}$$

By comparing this with the general relationship between incident and reflected radiances, it can be seen that Eq. (2-61) will result if the partial reflectance ρ' in Eq. (2-47) has the form

$$\rho' = 2\rho_d(\theta_i, \varphi_i)\delta(\sin^2\theta_r - \sin^2\theta_i)\delta(\varphi_r - \varphi_i \pm \pi) \qquad \text{sr}^{-1} \tag{2-62}$$

where $\delta(\sin^2\theta_r - \sin^2\theta_i)$ and $\delta(\varphi_r - \varphi_i \pm \pi)$ are Dirac delta functions which satisfy the defining relations

$$\delta(u) = 0 \qquad \text{for } u \neq 0$$

$$\int \delta(u)\, du = 1$$

and

$$\int f(u)\delta(u)\, du = f(0)$$

when the integration is carried out over the full range of the variable, $0 \leqslant \theta \leqslant \pi/2$ and $0 \leqslant \varphi \leqslant 2\pi$, in each case.

Only what might be termed the external radiometric relations have been considered in the foregoing treatment, and no attempt has been made to deal with the deeper theory relating reflectance, emissivity, and absorptance to the optical constants of the materials. A good summary of the most important aspects of that approach is given in [33].

2.11. Summary of Equations and Constants

$W = \sigma T^4$ (2-21)

$W_\lambda = 2\pi c^2 h \lambda^{-5}(e^{hc/kT} - 1)^{-1}$ (2-8)

$W_\nu = 2\pi c^{-2} h \nu^3 (e^{h\nu/kT} - 1)^{-1}$ [follows from (2-8) and (2-11)]

$N_\lambda = W_\lambda/\pi$ (2-19) (only Lambertian surfaces, *e.g.*, blackbody)

$U\lambda = 1.2398 \text{ ev } \mu$ (2-14)

$u_\lambda = 4W_\lambda/c$ (2-10*a*) (within isothermal enclosure)

$u_\lambda = 2W_\lambda/c$ (2-10*b*) (into Lambertian hemisphere)

$n_\nu = W_\nu/h\nu = 2\pi c^{-2}\nu^2(e^{h\nu/kT} - 1)^{-1}$ (2-15)

$n_\lambda = W_\lambda/hc = 2\pi c\lambda^{-4}(e^{hc/\lambda kT} - 1)^{-1}$ (2-16)

$W_\lambda\, d\lambda = W_\nu\, d\nu$ (2-12)

$\overline{(\Delta n)^2} = n[1 + (e^{h\nu/kT} - 1)^{-1}] \approx n$ (2-17) and (2-18)

$\lambda_{max} T = 2897.9\ (\mu\,^\circ\text{K})$ (2-24)

$$\lambda_{max\ diff.} T = 2404\ (\mu\ °\mathrm{K}) \qquad (2\text{-}26)$$

$$c = \nu\lambda$$

$$\sigma = \tilde{\nu} = k/2\pi = 1/\lambda\ \text{(wavenumber)}$$

$$d\nu = -c\ d\lambda/\lambda^2$$

$$d\lambda = -c\ d\nu/\nu^2$$

$$d\lambda/\lambda = d\nu/\nu$$

$$J = P_\Omega = \partial P/\partial\Omega$$

$$\left.\begin{aligned} W &= P_A \\ H &= P_A \end{aligned}\right\} = \partial P/\partial A$$

$$N = P_{A\Omega} = \frac{\partial^2 P}{\cos\theta\ \partial A\ \partial\Omega} = \frac{\partial^2 P}{\partial A\ \partial\Omega'} \text{ where } d\Omega' = \cos\theta\ d\Omega$$

$$h = 6.6252 \times 10^{-34}\ \text{w sec}^2$$

$$\pi = 3.1416$$

$$c = 2.99793 \times 10^8\ \text{m sec}^{-1}$$

$$\sigma = \frac{2\pi^5 k^4}{15c^2h^3} = 5.6686 \times 10^{-12}\ \text{watt cm}^{-2}\ (°\mathrm{K})^{-4}\ \text{(Stefan-Boltzmann constant)}$$

$$k = 1.38047 \times 10^{-16}\ \text{erg}\ (°\mathrm{K})^{-1}\ \text{(Boltzmann's constant)}$$

$$e = 2.71828\ \text{(base of Napierian logarithms)}$$

$$h/k = 4.079 \times 10^{-11}\ \text{sec}\ °\mathrm{K}$$

Quantity	c_1	c_1
Energy density	$8\pi ch$	4.99 joule $\mu^4\ \mathrm{m}^{-3}$
Emittance	$2\pi c^2 h$	$3.7413 \times 10^8\ \mathrm{w}\ \mu^4\ \mathrm{m}^{-2}$
Radiance	$2c^2 h$	$1.19 \times 10^8\ \mathrm{w}\ \mu^4\ \mathrm{m}^{-2}\ \mathrm{sr}^{-1}$

$$c_2 = ch/k = 1.4388 \times 10^4\ \mu\ °\mathrm{K}$$

References

1. M. Holter *et al., Fundamentals of Infrared Technology,* Chap. 1, Macmillan, New York (1963).
2. *Report of the Working Group on Infrared Backgrounds,* Part II:"Concepts and Units for the Presentation of Infrared Background Information," Report No. 2389-3-S, The University of Michigan, Institute of Science and Technology, Ann Arbor, Mich. (1956) AD 123 097.
3. S. S. Penner, *Quantitative Molecular Spectroscopy and Gas Emissivities,* Addison-Wesley, Reading, Mass. (1959).
4. G. Worthing and D. Halliday, *Heat,* Wiley, New York, 435 (1948).
5. F. E. Nicodemus, "Radiance," *Am. J. Phys.,* **31**, 5, 368 (1963).
6. A. H. Canada, *Gen. Elec. Rev.,* **51**, 50-54 (1948).
7. M. W. Makowski, *Rev. Sci. Inst.,* **20**, 876 (1945).
8. T. P. Merritt and F. F. Hall, Jr., *Proc. IRE* Sept. (1959).
9. S. A. Golden, *Spectral and Integrated Blackbody Radiation Functions,* Research Report 60-23, Rocketdyne Division, North American Aviation, Inc., Canoga Park, Calif. (1960).
10. M. Pivovonsky and M. Nagel, *Tables of Blackbody Radiation Functions,* Macmillan, New York (1961).
11. M. Czerny and A. Walther, *Tables of the Fractional Function for the Planck Radiation Law,* Springer-Verlag, Berlin (1961).
12. Parry Moon, *J. Opt. Soc. Am.,* **38**, 291 (1948).
13. A. N. Lowen and G. Blanch, *J. Opt. Soc. Am.,* **30**, 70 (1940).
14. E. Jahnke and F. Emde, *Tables of Functions,* Dover, New York (1945).

15. A. G. DeBell, Rocketdyne Research Report 59-32, Rocketdyne Division, North American Aviation, Inc., Canoga Park, Calif. (1959).
16. Henry Blau and Heinz Fischer (eds.), *Radiative Transfer from Solid Materials,* Macmillan, New York (1962).
17. R. Clark Jones, "Immersed Radiation Detectors," *Appl. Opt.,* **1,** 607 (1962).
18. D. K. Edwards, J. T. Gier, K. E. Nelson, and R. D. Roddick, "Integrating Sphere for Imperfectly Diffuse Samples," *J. Opt. Soc. Am.,* **51,** 1279 (1961).
19. D. E. Kerr, "Application of the Lorentz Reciprocity Theorem to Scattering" Appendix A (p. 693) of *Propagation of Short Radio Waves,* Vol. 13 of M.I.T. Radiation Laboratory Series (McGraw-Hill, New York, 1951), first edition.
20. A. T. DeHoop "A Reciprocity Theorem for the Electromagnetic Field Scattered by an Obstacle," *Appl. Sci. Res.,* Sec. B, **8,** 135 (1960).
21. H. J. McNicholas, "Absolute Methods in Reflectometry," *Bu. Std. J. Res.,* I, **29** (RP-3), submitted to The Johns Hopkins University, PhD, dissertation (1928).
22. J. C. DeVos, "Evaluation of the Quality of a Blackbody," *Physica,* **20,** 669 (1954).
23. Max Born and Emil Wolf, *Principles of Optics,* **380**, Pergamon, New York (1959).
24. Committee on Colorimetry, Optical Society of America, "The Science of Color," 178, Crowell, New York (1954).
25. H. von Helmholtz, "Helmholtz's Treatise on Physiological Optics," James P. C. Southall (ed.) translated from Third German Edition, Optical Society of America, I, **231,** (1924).
26. M. Planck, *Theory of Heat,* translated by H. L. Brose, **5,** 198, Macmillan, New York (1957).
27. Bruce Hapke and Hugh Van Horn, "Photometric Studies of Complex Surfaces, with Applications to the Moon," *J. Geophys. Res.* **68,** 15, 4545 (Aug. 1, 1963).
28. Bruce W. Hapke, "A Theoretical Photometric Function for the Lunar Surface," *J. Geophys. Res.* **68,** 15, 4571 (Aug. 1, 1963).
29. V. Twersky, *IRE Trans. Antennas and Propagation,* **AP-5**, 81 (1957).
30. V. Twersky, *J. Res. NBS,* **64D**, 715 (1960).
31. V. Twersky, "Multiple Scattering of Waves and Optical Phenomena" *J. Opt. Soc. Am.,* **52,** 145 (1962).
32. H. E. Bennett and J. O. Porteus, "Relation Between Surface Roughness and Specular Reflectance at Normal Incidence," *J. Opt. Soc. Am.,* **51,** 123 (Feb. 1961).
33. Henry H. Blau, Jr., John L. Miles, and Leland E. Ashman, *Thermal Radiation Characteristics of Solid Materials – a Review,* Arthur D. Little, Inc. Contract No. AF 19 (604)-2639, Scientific Report No. 1, AFCRC-TN-58-132 (1958) AD 146 883.

Chapter 3

ARTIFICIAL SOURCES*

CONTENTS

3.1. Theory 32
3.1.1. The Method of Gouffé 32
3.1.2. The Method of DeVos 33
3.1.3. The Method of Sparrow. 37
3.2. NBS Standards 38
3.2.1. Instructions for Using the Total Radiation Standards. 38
3.2.2. Instructions for Using the NBS Standards of Spectral Irradiance . . 40
3.2.3. Instructions for Using Standards of Spectral Radiance 41
3.2.4. Other NBS Sources 44
3.3. Laboratory and Field Sources 46
3.3.1. A High-Temperature Source 46
3.3.2. Low-Temperature Large-Area Sources. 47
3.3.3. Nernst Glower. 48
3.3.4. Globar 48
3.3.5. Welsbach Mantle 48
3.3.6. Carbon Arc 49
3.3.7. Tungsten Filaments 49
3.3.8. Mercury Arcs 50
3.3.9. Zirconium Point Sources 51
3.4. Commercial Cavity-Type Sources 51

*Material prepared by the technical writing staff of McGraw-Hill, Inc.

3. Artificial Sources

3.1. Theory

The classic blackbody simulator is a spherical cavity made of an opaque material and pierced by a small aperture. If the enclosure is insulated from outside thermal influences, all parts of the internal cavity walls eventually reach the same temperature. When equilibrium is reached, all points inside the cavity have equal energy regardless of the nature of the walls, and the radiation passing out of the small aperture approximates blackbody radiation; the smaller the hole the better the approximation.

A practical cavity approaches a blackbody to the extent that the radiation entering through the aperture is absorbed within the cavity. Since all practical materials have absorptivities less than unity, the performance of a blackbody-simulating cavity is based on the number of reflections that radiation entering the cavity through the aperture can make before returning through the aperture to the outside.

DeVos [1], Edwards [2], Gouffé [3], Williams [4], Sparrow [5], and Zissis [6] have written theoretical analyses of cavity radiation. Williams' paper contains a criticism of the others. The most thorough treatment seems to be that of DeVos, although Williams' criticisms of it are valid. Truenfels [7] has also written a pertinent article. Three techniques are given below (DeVos, Gouffé, and Sparrow). No direct detailed comparison of them is available.

3.1.1. The Method of Gouffé. For the total emissivity of the cavity forming a blackbody, Gouffé gives:

$$\epsilon_0 = \epsilon'_0 (1 + k) \tag{3-1}$$

where

$$\epsilon'_0 = \frac{\epsilon}{\epsilon[1 - (s/S)] + (s/S)} \tag{3-2}$$

and $k = (1 - \epsilon)\,[(s/S) - (s/S_0)]$, and is always nearly zero; it can be either positive or negative

ϵ = emissivity of materials forming the blackbody surface

s = area of aperture

S = area of interior surface

S_0 = the surface of a sphere of the same depth as the cavity in the direction normal to the aperture

Figure 3-1 is a graph for determining the emissivities of cavities with simple geometric shapes. (Values not legible can be calculated without great difficulty.) In the lower section, the value of the ratio s/S as a function of the ratio L/R is read (L and R are defined in Fig. 3-1). Reading up from this value to the value of the intrinsic emissivity of the cavity material, the value of ϵ'_0 is found. Multiplying ϵ'_0 by the factor $(1 + k)$ then yields the emissivity of the cavity. Values of s/S, etc., are given in Table 3-1.

When the aperture diameter is smaller than the interior diameter of the cylindrical cavity or the base diameter of the cone forming a conical cavity, it is necessary to multiply the values of s/S determined from the graph by $(r/R)^2$, which is the ratio of the squares of the aperture and cavity radii (Fig. 3-1).

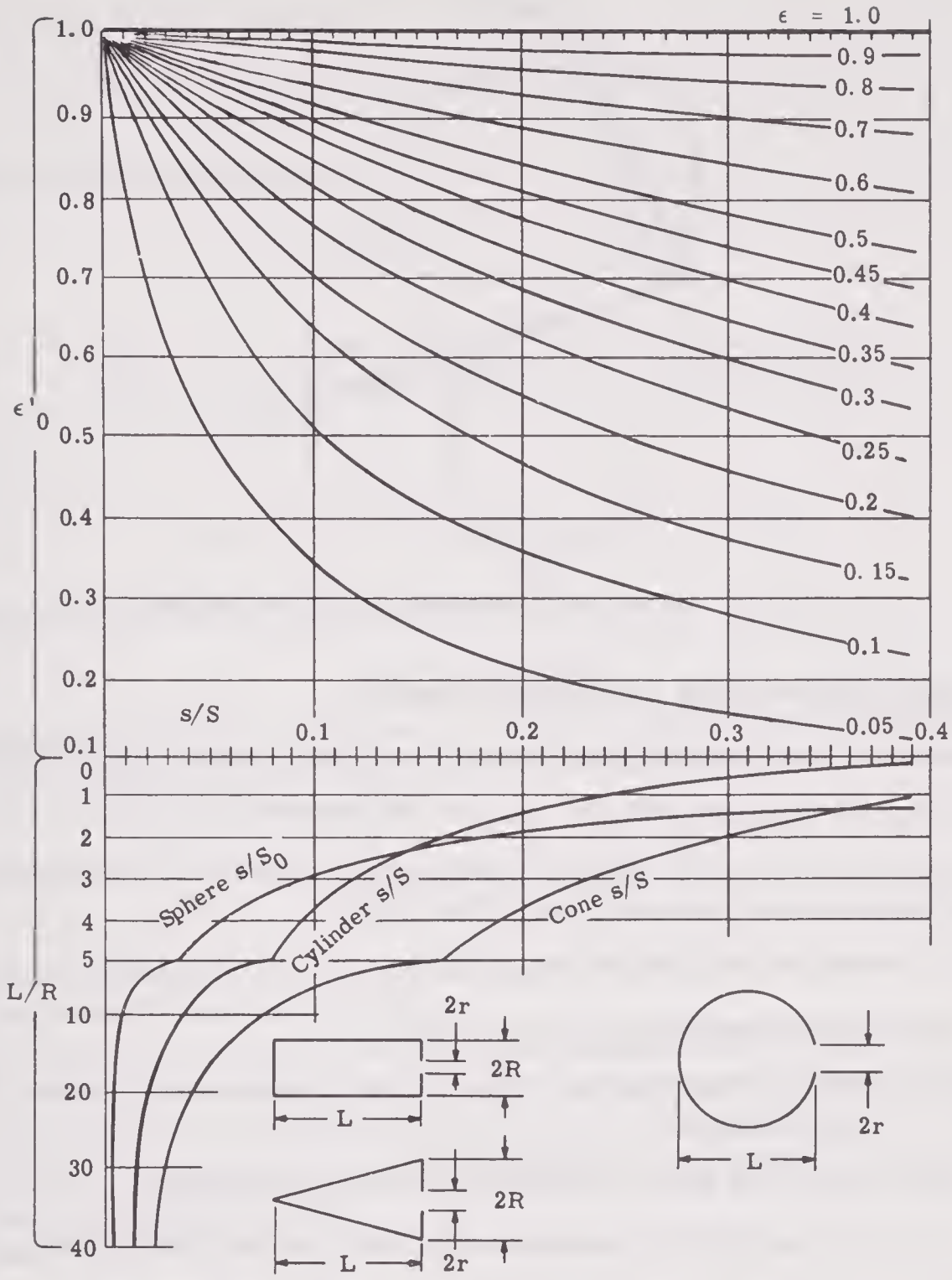


FIG. 3-1. Emissivities of conical spherical and cylindrical cavities.

3.1.2. The Method of DeVos. DeVos considers a cavity of arbitrary shape, with opaque walls, in a nonattenuating medium, initially at a uniform steady temperature, with one small opening. He adds additional openings and temperature variations along the cavity walls, and indicates several practical approximations which are necessary for calculation of numerical values. See Fig. 3-2 for definitions of terms.

The power emitted from $d0$ is

$$^{em}P_{w}{}^{0} = \epsilon_{w}{}^{0}(\lambda, T)N_{\lambda,B}(\lambda, T)\, dw \cos \theta_{w}{}^{0}\, d\Omega_{w}{}^{0} \tag{3-3}$$

where $\epsilon_{w}{}^{0}(\lambda, T)$ = the spectral emissivity of dw in the direction of $d0$ (indicated by sub- and superscripts throughout) for temperature T and at wavelength λ

$N_{\lambda,B}(\lambda, T)$ = the spectral radiance of a blackbody for temperature T and wavelength λ, given by either modification of Eq. (3-2) or approximately by the Wien law,

$$N_{\lambda,B} = (\text{constant})\, e^{-c_2/\lambda T} \tag{3-4}$$

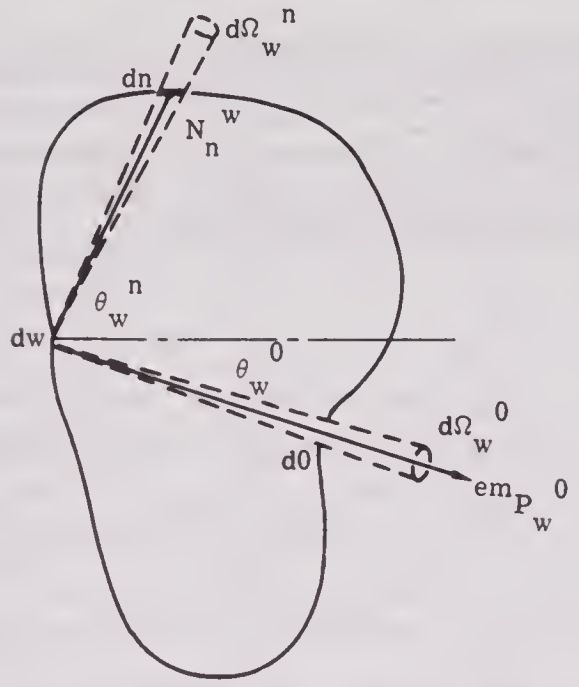


FIG. 3-2. Definition of terms for the DeVos method.

dw = area of the emitting infinitesimal element

$\theta_w{}^0$ = the angle of the direction from dw to $d0$ with respect to the normal to dw

$d\Omega_w{}^0$ = the solid angle subtended by $d0$, the hole, as seen from dw

The power from dw through $d0$ which is due to the reflection of the power received at dw from some arbitrary elemental wall area dn is

$$^{refl}P_w{}^{n0} = N_{\lambda,n}{}^w(\lambda, T)\, d\Omega_w{}^n\, dw \cos\theta_w{}^n\, r_w{}^{n0}(\lambda, T)\, d\Omega_w{}^0 \tag{3-5}$$

where $P_w{}^{n0}$ = the power from dn to $d0$ via dw

$N_{\lambda,n}{}^w(\lambda, T)$ = the spectral radiance from dn to dw for a temperature T and at wavelength λ

$d\Omega_w{}^n$ = the solid angle subtended by dn as seen from dw

$\theta_w{}^n$ = the angle to the normal to dw made by the direction from dw to dn

$r_w{}^{n0}(\lambda, T)$ = the partial reflectivity of dw for radiation from dn at $\theta_w{}^n$ reflected from dw toward $d0$ at $\theta_w{}^0$, at wavelength λ and for temperature T

Partial reflectivity can be defined as follows:

$$r_w{}^{n0} = \frac{N_w{}^{n0} \cos\theta_w{}^0}{N_n{}^w \cos\theta_w{}^n\, d\Omega_w{}^n} \tag{3-6}$$

As shown in Fig. 3-2, the reflected rays are not necessarily in the same plane. However, symmetry of the partial reflectivity about the plane containing the normal and the incident direction is assumed.

$$r_w{}^{n0} = \frac{N_w{}^{n0}}{H_w{}^n} \cos\theta_w{}^0 \tag{3-7}$$

To obtain the power from dw due to the reflection from dw of the radiation from all parts of the cavity walls except $d0$, integrate Eq. (3-5) over the walls excluding $d0$:

$$^{refl}P_w{}^0 = dw\, d\Omega_w{}^0 \int_{all\ dn} N_{\lambda,n}{}^w(\lambda, T) \cos\theta_w{}^n r_w{}^{0n}(\lambda, T)\, d\Omega_w{}^n \tag{3-8}$$

The reciprocity relation is:

$$r^{ab} \cos\theta^a = r^{ba} \cos\theta^b \tag{3-9}$$

By using the reciprocity relation in Eq. (3-9):

$$^{refl}P_{w}{}^{0} = dw\, d\Omega_{w}{}^{0} \cos\theta_{w}{}^{0} \int N_{\lambda,n}{}^{w}(\lambda, T) r_{w}{}^{0n}(\lambda, T)\, d\Omega_{w}{}^{n} \tag{3-10}$$

To a first-order approximation the following relationship is true:

$$N_{\lambda,n}{}^{w}(\lambda, T) = N_{\lambda,B}(\lambda, T)$$

Thus

$$\int_{all\ dn} r_{w}{}^{0n}(\lambda, T)\, d\Omega_{w}{}^{n} = P_{w}{}^{0}(\lambda, T) - r_{w}{}^{\infty}(\lambda, T)\, d\Omega_{w}{}^{0} \tag{3-11}$$

Therefore

$$^{total}P_{w}{}^{0} = N_{\lambda,B}\, dw \cos\theta_{w}{}^{0}\, d\Omega_{w}{}^{0}(1 - r_{w}{}^{\infty}\, d\Omega_{w}{}^{0}) \tag{3-12}$$

The hole can be considered to have an emissivity given by

$$\epsilon_{0} = 1 - r_{w}{}^{\infty}\, d\Omega_{w}{}^{0} \tag{3-13}$$

to a first approximation. If additional holes exist, or an opening like a slit which should be considered as several holes, then the reflected contributions of these elements must also be excluded. This leads to:

$$\epsilon_{0} = 1 - \sum_{h} r_{w}{}^{0h}\, d\Omega_{w}{}^{h} \tag{3-14}$$

for the emissivity of $d0$ in the direction from dw when there exists several holes numbered from $h = 0$ to some finite integer.

For the second-order approximation DeVos considers the use of a value for $N_{\lambda,n}{}^{w}$ which is not $N_{\lambda,B}$, but which is calculated by considering the effects of the holes on this spectral radiance (from each element dn) and the effect of $T_{n} \neq T_{w}$. The calculation neglects variation, due to temperature differences, of the emissivities and reflectivities of the wall elements.

The second-order approximation for the quantity $^{total}P_{w}{}^{0}$ is

$$^{total}P_{w}{}^{0} = N_{\lambda,B}(\lambda, T)\, dw \cos\theta_{w}{}^{0}\, d\Omega_{w}{}^{0} \left(1 - \sum_{h} r_{w}{}^{0h}\, d\Omega_{w}{}^{h} - \sum_{h} \int r_{n}{}^{wh}\right.$$

$$\left. d\Omega_{n}{}^{h} r_{w}{}^{0n}\, d\Omega_{w}{}^{n} - \int k_{n}\epsilon_{n}{}^{w} r_{w}{}^{0n}\, d\Omega_{w}{}^{n}\right) \tag{3-15}$$

where the integration is over the entire surface excluding the holes.

DeVos uses Wien's equation to evaluate k_{n} for cases where $T - T_{n} << T$. Thus,

$$N_{\lambda,B} \sim e^{-c_{2}/\lambda T} \tag{3-16}$$

and

$$\frac{dN_{\lambda,B}}{N_{\lambda,B}} = -\frac{c_{2}}{\lambda T}\frac{dT}{T} \tag{3-17}$$

where the constant $c_{2} = hc/k$ for T in °K and λ in cm. Thus

$$k_{n} \simeq \frac{c_{2}}{\lambda T}\frac{\Delta T}{T} \tag{3-18}$$

DeVos applied this theory to the V-wedge, cylindrical (closed at one end), spherical, and tubular shapes. He did not treat the cone, a combination of cones, or a cone-cylinder. He included temperature gradients only in the tubular calculation.

For a cylinder of radius R and length L (see Fig. 3-1), the value of ϵ_0 to first-order approximation is:

$$\epsilon_0 \simeq \left(1 - R_w{}^{\infty} \frac{\pi r^2}{L^2}\right) = \left(1 - R_w{}^{\infty} \frac{\pi}{\alpha^2}\right) \tag{3-19}$$

where α is the cone angle.

For the second-order approximation (neglecting temperature gradients), one needs $d\Omega_w{}^n$ and $d\Omega_w{}^0$. If dn is an annulus of the cylinder with a length dx, then

$$d\Omega_w{}^n = d\Omega_w{}^x = \frac{2\pi r^2\, dx}{[(L-x)^2 + r^2]^{3/2}} \tag{3-20}$$

and

$$d\Omega_x{}^0 = \frac{\pi r^2}{x^2 + r^2} \tag{3-21}$$

if $x/r = y$, then

$$\epsilon_0 = 1 - r_w{}^{\infty} \frac{\pi}{\alpha^2} - 2\pi^2 \int_0^{\alpha} \frac{r_w{}^0{}_y r_y{}^{w0}}{(y^2+1)\,[(y-\alpha)^2+1]^{3/2}}\, dy \tag{3-22}$$

DeVos evaluated this expression by numerical integration. His values, corrected by Edwards [2] for a small numerical error, are given in Table 3-1. For a similar calculation for a sphere, DeVos obtained the results in Table 3-2.

TABLE 3-1. DEVOS' EMISSIVITIES, CYLINDRICAL BLACKBODY [1]

(Emissivity values for a cylindrical blackbody with second-order corrections for various values of a ($= D/r$ = depth of cylinder/radius of cylinder) and surfaces of different smoothness. These are DeVos' values corrected for a numerical error. DeVos' figure numbers head the columns.)

a	*Fig.* 2	*Fig.* 3	*Fig.* 4	*Fig.* 5
6	0.970	0.954	0.865	0.668
10	0.990	0.985	0.953	0.864
15	0.995	0.994	0.980	0.947
20	0.997	0.997	0.989	0.972
30	0.999	0.999	0.996	0.988

TABLE 3-2. DEVOS' EMISSIVITIES, SPHERICAL BLACKBODY [1]

(Emissivity values for a spherical blackbody with second-order corrections for various values of a ($= D/r$ = diameter of sphere/radius of hole) and surfaces of different smoothness. These are DeVos' values corrected for a numerical error. DeVos' figure numbers head the the columns.)

a	*Fig.* 2	*Fig.* 3	*Fig.* 4	*Fig.* 5
10	0.992	0.989	0.963	0.894
20	0.998	0.998	0.991	0.976

3.1.3. The Method of Sparrow [8]. "The starting point of the analysis is a radiant flux balance at a typical position x_0 on the cylindrical wall. Such a balance states that the radiant energy leaving a location is equal to the emission plus the reflected incident energy.

$$\epsilon_a(x'_0) = \epsilon + (1-\epsilon)\int_0^{L/d} \epsilon_a(x') \left\{1 - |x'_0 - x'| \frac{2(x'-x'_0)^2+3}{2[(x'-x'_0)^2+1]^{3/2}}\right\} dx'$$

$$+ 4(1-\epsilon)\left(\frac{L}{d} - x'_0\right)\int_0^1 \epsilon_a(r') \frac{\left\{4\left(\frac{L}{d}-x'_0\right)^2 + 1 - r'^2\right\} r'}{\left\{\left[4\left(\frac{L}{d}-x'_0\right)^2 + 1 + r'^2\right]^2 - 4r'^2\right\}^{3/2}} dr' \qquad (3\text{-}23)$$

$$\epsilon_a(r') =$$

$$\epsilon + 8(1-\epsilon)\int_0^{L/d} \epsilon_a(x')\left(\frac{L}{d} - x'\right) \frac{4\left(\frac{L}{d}-x'\right)^2 + 1 - r'^2}{\left\{\left[4\left(\frac{L}{d}-x'\right)^2 + 1 + r'^2\right]^2 - 4r'^2\right\}^{3/2}} dx' \qquad (3\text{-}24)$$

where $\epsilon_a = N/\sigma T^4$, $x' = x/d$, $r' = r/R$.

"It may be seen that $\epsilon_a(x')$ and $\epsilon_a(r')$ appear in both Eqs. (3-23) and (3-24), which therefore require simultaneous solution. Additionally, since the unknowns appear under integral signs, these equations are integral equations. Further inspection reveals that there are two parameters which may be independently prescribed: L/d, which describes the relative length of the hole; and ϵ, which characterizes the radiation properties of the materials.

"The mathematical problem as represented by these equations is complex, but this is the price which must be paid in order to achieve a complete formulation of the physical occurrences within the radiating enclosure... . By assuming $\epsilon_a(r') =$ constant, the system was reduced from two simultaneous integral equations to a single integral equation. Further, all angle factors were written approximately in terms of exponentials of the form $e^{-|x_0 - x|}$, etc., rather than as the precise, but more complex, expressions which appear in Eqs. (3-23) and (3-24). The effect of these simplifications has heretofore remained unevaluated for finite-length cavities.

"The formidable mathematical problem here is not amenable to a closed-form solution, and it was necessary to use numerical means. The calculation scheme was essentially a direct iteration, the steps of which may be outlined as follows: First, values of the parameters L/d and ϵ were chosen in order to specify the problem. Next, a trial solution for $\epsilon_a(x')$ was guessed over the range $0 \leq x' \leq L/d$. As will be described later, it was possible to make very favorable guesses by taking advantage of prior solutions. These guessed values for $\epsilon_a(x')$ served as input data to the right side of Eq. (3-24). For a given value of radial position r', the integration was carried out numerically to yield the value of ϵ_a at that r'. This could be repeated for each point in the range $0 \leq r' \leq 1$, and in this way there was generated a function $\epsilon_a(r')$ corresponding to the guessed distribution $\epsilon_a(x')$. The $\epsilon_a(r')$ and $\epsilon_a(x')$ were then utilized in evaluating the integrals of Eq. (3-23). For a fixed x'_0, numerical integration of Eq. (3-23) provided an ϵ_a corresponding to that x'_0. By repeating the process for all x'_0 in the range $0 \leq x'_0 \leq L/d$, a new function $\epsilon_a(x')$ or $\epsilon_a(x'_0)$ was obtained which could be used as input data to Eq. (3-24), thereby beginning another cycle in the iteration. This procedure led to a steady convergence to the final result, and stability problems were not encountered.

"After a few solutions were carried out, it was found that, for a fixed ϵ, the distribution of ϵ_a along the cylindrical wall was not highly sensitive to the L/d ratio. Thus the $L/d = \infty$ results from Reference [9] were utilized as the first guess for $\epsilon_a(x)$ for the $L/d = 4$ case; the converged results for $L/d = 4$ were utilized for starting values for $\epsilon_a(x)$ for the $L/d = 3$ case; etc. This ordering of the solutions – from high to low L/d values – helped somewhat to ease what still remained a formidable computational task.

"The actual numerical calculations were performed on an IBM 704 electronic digital computer. Solutions were carried out for $L/d = 4$, 3, 2, 1, 0.5, and 0.25 for $\epsilon = 0.9$, 0.75, and 0.5. For this emissivity range, it was found unnecessary to extend the calculations to lengths beyond $L/d = 4$, since nearly blackbody conditions already prevailed over the end disk."

The article presents results only for cylindrical cavities, shapes which are not of the greatest practical importance.

3.2. NBS Standards

The National Bureau of Standards does not have an official standard blackbody which it uses for comparison tests. It does, however, check thermocouples and other temperature-measuring devices. As the section on theory (Sec. 3.1.3) shows, it is easy to design a blackbody which gives high emissivity, but the temperature must be known accurately. The Bureau does prepare and issue secondary standards in the form of total radiance and irradiance lamps. Instructions for the use of these sources are given below:

3.2.1. "Instructions for Using the Total Radiation Standards (Revised February 24, 1960)"

"These instructions cover the use of our small carbon filament standards of thermal radiation. The lamps employed (about 50 watts) have been seasoned, marked for orientation, and calibrated for density of radiant flux at a fixed distance in a specified direction from the lamp. Suitable markings (a line on one side and a line through a circle on the opposite side) have been etched into the glass of the lamp bulbs to assist the operator in setting up the lamps relative to the radiometer in a position identical to that under which the standards were calibrated.

"In operation the lamp is to be screwed into an ordinary metal socket that is held upright by a support, which cannot reflect light into the radiometer. The lamp is to be oriented so that the etched line on one side and the line and circle on the opposite side are in line with the radiometer. The circle on the lamp is to be situated away from the radiometer and on the horizontal axis through the radiometer. The entire lamp bulb is to be exposed to the radiometer. Sufficient time must be given (say five minutes) for the glass base, which supports the filament, to become thoroughly warmed, otherwise errors will be introduced into the radiant energy measurements. The distance of the lamp is measured from the radiometer to the center (glass tip if present, the etched mark, or other special mark noted in the 'Report' on a particular standard lamp) of the lamp bulb.

"A black cloth, of about 1 meter by 1 meter edge, should be placed about 1 meter to the rear of the lamp. An opaque shield about 1 m by 1 m, having an opening about 10 cm wide and 15 cm high, is placed at a distance of about 25 cm in front of the lamp. To screen this opening, a shutter, about 20 by 20 cm, is placed between this shield and the lamp. Facing the opening in the shield, the radiometer is placed at a distance of 2 m from the lamp. In this manner constant extraneous

conditions are maintained between the lamp and the radiometer when the shutter is opened and closed. The shield and shutter may be made of air-separated sheets of cardboard, asbestos board, metal or other suitable materials.

"Before the lamp is lighted, the shutter should be opened and closed to determine the amount of stray thermal radiation falling upon the radiometer. This test may be applied at any time provided the lamp has been given sufficient time to come to room temperature. The wall and screen to the rear of the lamp may be cooler than the shutter, which will cause a negative deflection. The correction to the observed lamp deflection is, in that case, positive. It is desirable to make the calibration in a dimly lighted room to avoid errors from sunlight which is continually varying with cloudiness, thus varying stray radiant energy within the room as well as the temperature of the walls, and also causing air currents near the radiometer.

"The values of radiant flux of this Bureau's primary standard of radiation are based upon direct measurements in absolute value; and upon a direct comparison with a blackbody using the Stefan-Boltzmann constant of total radiation of a blackbody $\delta = 5.7 \times 10^{-12}$ watt per cm^2, as described in *NBS Bulletin* No. 227, Vol. 11, p. 87, 1914 [10], and *NBS, Journal of Research* No. 578, Vol. 11, p. 79, 1933 [11]. The absolute values of the primary standard are accurate to about 1 percent. The values of the secondary standards, compared with the primary standard, are in agreement within about 0.5 percent. The overall accuracy in the use of these standards is somewhat dependent upon the conditions of temperature and humidity existing during their operation. Highest accuracy will appear for a room temperature of about 25°C and a relative humidity of about 60 percent near which the original calibrations were made. For extreme conditions of temperature and of relative humidity, corrections may be required, but can usually be neglected (see *NBS Journal of Research*, Vol. 53, p. 211, 1954) [12].

"The best results are to be obtained by operating the lamp on 0.30 to 0.35 ampere. This is the current through the lamp. If the measurement of current be made when a voltmeter is in the circuit with the lamp, then a correction may have to be made to the observed current.

"The measurement of current through the lamp is, of course, sufficient to determine the radiant flux, the voltage being useful mainly to determine whether the lamp has remained constant.

"To conserve the calibration, which gradually changes with use, these lamps should be kept as reference standards only, and other lamps used as working standards in all cases where extensive radiometric comparisons are made.

"These instructions and standards of radiation apply to radiometers used in air. If a window be used on the radiometer, as for example in a vacuum radiometer, then a correction has to be made for the radiant flux absorbed by the window, for the particular lamp used as a standard and for the source measured. This absorption is a function of the temperature of the lamp filament. For example, it was found that for a glass or quartz window about 1.5 mm in thickness, the transmission amounted to about 83 percent when a certain standard lamp was operated on 0.35 amp., and increased to 84 percent when the lamp was operated on 0.40 amp. Using a fluorite window, the transmission is higher (about 91.5 percent) and varies but little with the current ordinarily used in the lamp. For example, using a certain standard lamp, the transmission through a fluorite window varied from 91.0 percent on 0.25 amp. to 92 percent on 0.4 amp., with an average value of 91.6 percent on 0.35 amp.

"The transmission of the window varies also with the spectral quality of the radiant flux emitted by the source under investigation. This must also be taken into consideration.

"The thermal radiation sensitivity of a surface thermopile varies with the degree of evacuation; when highly evacuated this sensitivity may be several times as great as in air. Since at low air pressures the sensitivity is variable with the pressure, great care must be taken to test the thermopiles sensitivity under the exact conditions existing during its use.

"The same area of the radiometer receiver should be exposed to the standard of radiation as is used in the measurements of the unknown source."

3.2.2. "Instructions for Using the NBS Standards of Spectral Irradiance (May 31, 1963)"

"These instructions cover the use of tungsten-filament quartz-iodine lamps issued as standards of spectral irradiance for the wavelength range of 0.25 to 2.6 microns. The lamps employed are commercial G.E. type 6.6A/T4Q/1CL-200-watt lamps having a tungsten coiled-coil filament enclosed in a small (1/2 inch × 2 inches) quartz envelope containing a small amount of iodine.

"The spectral radiant intensity of the entire lamp as mounted in the manner prescribed below is measured and reported. The spectral irradiance from these lamps is based upon the spectral radiance of a blackbody as defined by Planck's equation and has been determined through comparison of a group of quartz iodine lamps with (1) the NBS standards of spectral radiance, (2) the NBS standards of luminous intensity, and (3) the NBS standards of total irradiance.

"The lamp is mounted vertically with the NBS-numbered end of the lamp down and the tip away from the detector. Measurements of distance (from lamp filament) are made along a horizontal axis passing through the center of the lamp filament. The correct vertical position is determined by setting the centers of the upper and lower seals along a plumb line as viewed from one side of the lamp. The plane of the front surface of the lower press seal is set to contain the horizontal perpendicular to the line connecting the lamp filament center and detector.

"The lamp is mounted in the supplied holder which is constructed in such a manner as to reflect a negligible amount of radiant flux in the direction of the radiometer or spectrometer slit. A black shield should be placed at a distance of about 3 feet to the rear of the lamp to intercept stray radiant flux along the radiometric axis and adequate shielding should be provided to intercept stray flux from other directions.

"If there is excessive water vapor in the laboratory atmosphere, errors may result at the wavelengths of water-vapor absorption bands. In the original calibrations the comparisons of the lamps with the other NBS standards were made in such a manner that the effect of water-vapor absorption was eliminated.

"Values of spectral irradiance for these lamps are tabulated as a function of wavelength in microwatts per (square centimeter-nanometer) at a distance of 43 centimeters from center of lamp to receiver. Values of spectral irradiance for wavelength intervals other than one nanometer, say x nanometers, may be found by multiplying the tabulated values by x.

Use of the Standards of Spectral Irradiance

"These standards require no auxiliary optics. If any are employed proper correction must be made for their optical characteristics. The lamp is simply placed at a measured distance from the detector or spectrometer slit. If a distance other

than 43 centimeters is used, the inverse-square law may be used to calculate the spectral irradiance (the inverse-square law should not, however, be used for distances shorter than about 40 centimeters).

"In measurements wherein two sources (a standard source and a test source) are being compared by the direct substitution method (slit widths kept unchanged, use of the same detector) no knowledge of the spectral transmittance of the spectrometer, nor of the spectral sensitivity of the detector is required. It is necessary, however, to make sure that the entrance slit of the spectrometer is fully and uniformly filled with radiant flux both from the standard and from the test source; and if at any one wavelength the detector response for the standard is significantly different from that for the test source, the deviation from linearity of response of the detector must be evaluated and taken into account. Furthermore, if the standard and test source differ in geometrical shape, it must be ascertained that the instrument transmittance and detector response are not adversely affected thereby. Many detectors are highly variable in sensitivity over their surface area and may require diffusion of radiant flux over their surface to insure accurate radiant energy evaluations.

"All calibrations were made by the use of alternating current and it is recommended that they be so used in service. To reduce the line voltage a 10-ampere variable autotransformer may be employed for coarse control. For fine control a second (5-ampere) variable autotransformer may be used to power a radio-filament transformer whose secondary (2.5-5 volt) winding is wired in series with the primary of the 10-ampere transformer. It was found that this method is very effective for accurate control of the 6.50-ampere current.

"These standards of spectral irradiance are expensive laboratory equipment and it is suggested that they be operated sparingly and with care in order to prolong their useful life. They should be turned on and off at reduced current and great care should be taken so that at no time will the current appreciably exceed 6.50 amperes. It is recommended that for general use, working standards be prepared by calibrating them relative to the laboratory standard supplied by NBS.

"These lamps operate at high temperature such that the quartz envelope is above the flammable point of organic materials. They may thus cause fires, and also the burning of lint, etc. on the enveloper which may result in optical damage to its surface. In no case should the fingers come into contact with the quartz envelope, either hot or cold, as the resulting fingerprints will burn into its surface during lamp operation.

3.2.3. "Instructions for Using Standards of Spectral Radiance (Revised February 21, 1961)"

General Discussion

"These instructions cover the use of ribbon-filament lamps issued as standards of spectral radiance for the wavelength ranges of 0.25 to 0.75, 0.5 to 2.5, and 0.25 to 2.5 microns. The lamps employed are commercial G.E. Type 30A/T24/7 lamps having a tungsten ribbon filament (SR-8A type) centered about 8 to 10 cm behind a fused silica window 3 cm in diameter.

"The portion of the filament whose spectral radiance has been determined is the central portion visible through the fused-silica window. This determination was made by direct substitution of the lamps for working standard lamps which had in turn been calibrated by this substitution method relative to blackbodies (operated from 1400 to 2300°K) through the use of a double quartz prism spectroradiometer and associated electronic equipment.

"In operation the lamp is mounted vertically and the beam of radiant flux with a horizontal axis passing through the center of the filament is measured. In the original determinations no portion of the beam measured departed from this axis by more than 2.5 degrees. Hence, if an aperture subtending a larger angle is required in any application of these standards of spectral radiance, it should be ascertained that the irradiance is constant over the whole aperture.

"If there is excessive water vapor in the laboratory atmosphere errors may result at the wavelengths of water-vapor absorption bands. In the original calibrations the comparisons of the lamps with the blackbodies were made at the same distance and in such manner that the effect of water-vapor absorption canceled out.

"In the calibrations wherein the blackbody was heated within a wirewound furnace and temperatures around 1400°K were reached, the temperatures were measured with Pt-Pt 10% Rh thermocouples. Tests using couples placed at various positions and observations with an optical pyrometer indicated closely uniform temperatures within the blackbody enclosure. A ratio of blackbody opening to total internal surface area equal to approximately 0.003 and an internal surface reflectance less than 0.10 indicate an emissivity of 0.999 or higher.

"For the blackbody temperatures above 1400°K a graphite enclosure heated by a radio-frequency generator was employed and the temperatures were measured by an optical pyrometer. The physical characteristics of this blackbody indicated an emissivity approximating 0.996.

"The spectral radiance of the blackbody is based upon the Planck radiation law in which the constants, based upon the most recent atomic and other information, are set down as follows:

$c_1 = 1.19088 \times 10^{-12}$ watt cm^2 per steradian

$c_2 = 1.4380$ cm degree K

"Values of spectral radiance for these lamps are tabulated as a function of wavelength in microwatts per (steradian-millimicron-square millimeter of filament). Values of spectral radiance for slit-widths other than one millimicron, say x millimicrons, where x is less than 100, may be found by multiplying the tabulated values by x.

Use of the Standards of Spectral Radiance

"It is suggested that the auxiliary optics employed with these standards be composed of two units: namely, a plane mirror and a spherical mirror (each aluminized on the front surface). If the spherical mirror is placed at a distance from the lamp filament equal to its radius of curvature, and the plane mirror set about 1/3 to 2/5 this distance from the spherical mirror, facing it and so placed that the angle between incident and reflected beams is 10° or less, a good image of the filament itself may be focused upon the spectrometer slit. Little distortion of the filament image occurs provided precise optical surfaces are employed and angles between incident and reflected beams are kept to less than 10°. Various optical arrangements may be employed.

"The solid-angular aperture of the auxiliary optics should be smaller than the solid-angular aperture of the spectrometer employed so that no loss of radiant energy will result through over-filling the spectrometer optics.

"The spectral radiant flux, P_λ, in microwatts per millimicron, which enters the spectrometer slit is computed from the formula:

$$P_\lambda = R_\lambda N_\lambda s A / D^2$$

where R_λ is the spectral reflectance of the combination of mirrors used, N_λ is the reported spectral radiance of the standard, s is the area of the spectrometer slit in mm^2, A is the area of the limiting auxiliary optic, and D is the distance of this optic from the slit.

"No diaphragm or other shielding is required in the use of these standards, except for a shield to prevent radiant energy from the lamp from entering the spectrometer directly without first falling on the concave mirror. An image of the filament should be focused upon the spectrometer slit, and only the energy by which this image is formed should enter the slit.

"In order to calibrate a spectroradiometer with one of these standards of spectral radiance, a knowledge of the spectral reflectance of the mirror surfaces is required. A good aluminized surface should have a spectral reflectance considerably above 0.87 throughout the spectral region of 0.5 to 2.6 microns, increasing slightly with wavelength except possibly for a slight dip near 0.80 micron. In practice the proper reflectance losses can best be determined through the use of a third (similar) mirror (a second plane mirror) which may be temporarily incorporated into the optical set-up from time to time.

"In measurements wherein two sources (a standard source and a test source) are being compared by the direct substitution method (use of the same auxiliary optics, slit-widths, areas and detector at any one wavelength), no knowledge of the spectral reflectance of the auxiliary mirrors, nor of the spectral transmittance of the spectrometer, nor of the spectral sensitivity of the detector is required. It is necessary, however, to make sure that the entrance slit of the spectrometer is fully and uniformly filled with radiant flux both from the standard and from the test source; and, if at any one wavelength the detector response for the standard is significantly different from that for the test source, the deviation from linearity of response of the detector must be evaluated and taken into account.

"Operation of these standards should be on alternating current to obviate filament-crystallizing effects that occur when the operation is on direct current. The filaments are massive and "iron out" all effects of the normal fluctuations present in a commercial ac supply. All calibrations were made by means of alternating current. To reduce line voltage a step-down transformer (1 kva) having a ratio of 10 to 1 or a 50-ampere variable transformer may be employed. Then to give fine control a second variable transformer (10-ampere capacity) is wired into the circuit to control the input of the heavy duty transformer. For still finer control a third variable transformer may be employed with a radio-filament transformer to add (or subtract) voltage fed into the step-down transformer. It was found that this method is very effective for accurate control of large lamp currents. The heavy duty (1 kva) step-down transformer is preferred to the 50-ampere variable transformer since the latter is subject to contact damage when operated for long intervals of time at high currents.

"The lamp standards of spectral radiance are expensive laboratory equipment and it is suggested that they be operated sparingly and with care in order to prolong their useful life. This precaution applies especially to the standards calibrated in the short-wave region and operated at 35 amperes. They should be turned on and off slowly and only for short intervals should they ever be operated at or above 30 amperes, and then only to calibrate a similar lamp as a working standard. In general even at lower currents a working standard should be prepared and used except for purposes of checking the operation of such a working standard."

3.2.4. Other NBS Sources. Figure 3-3 shows a blackbody furnace used at the National Bureau of Standards as a comparison standard in emissivity determinations [13]. The core, made of Inconel, has an outside diameter of 3 cm and a cylindrical cavity 2.06 cm in diameter and 7 cm deep with inner walls roughened by a very fine tap and treated to produce an opaque, oxidized layer. The furnace is heated by an 80% platinum-20% rhodium resistance element surrounding the core. The core and heating element are covered with thermal insulation 3.8 cm thick and are mounted axially in a water-cooled tube of 11.4 cm outside diameter. Power input to the furnace is adjusted manually by an autotransformer. The temperature of the furnace is measured by a thermocouple inserted in a hole drilled from the back of the core to within 0.08 cm of the inner surface of the core at the rear of the cavity. The emissivity of the inner surface is greater than 0.90 and that of the furnace was computed to be greater than 0.99 in the temperature range for which it was designed, 810° to 1360°K.

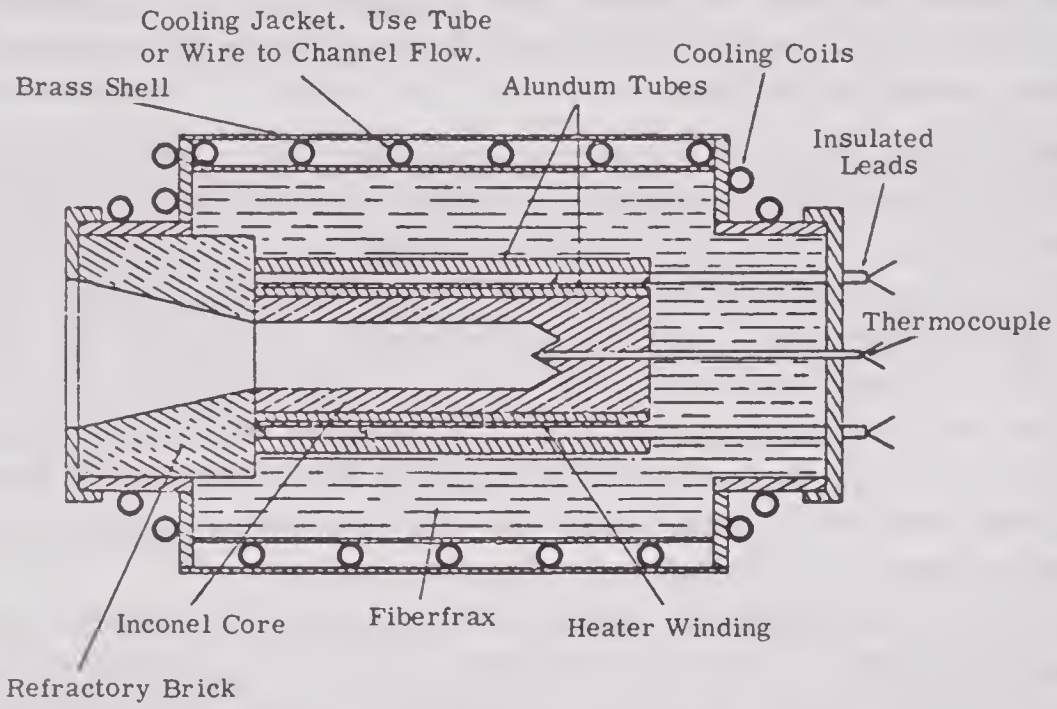


FIG. 3-3. NBS blackbody furnace.

Figures 3-4 and 3-5 show two other blackbody cavities designed by NBS. They are probably the closest approximation to blackbody radiators in existence. These cavities are primarily used in optical pyrometry as gold-point blackbodies to determine the International Practical Temperature Scale (IPTS) above 1336°K [14].

The vertical blackbody assembly shown in Fig. 3-4 is heated in a wire-wound muffle furnace or in the coils of an rf generator. The estimated emittance of this blackbody is 0.999, assuming the walls are at a uniform temperature. The horizontal blackbody shown in Fig. 3-5 has three independently controlled heater windings that are embedded longitudinally in cylindrical graphite muffles.

The graphite, having high thermal conductivity, tends to reduce longitudinal temperature gradients. The power inputs to the two end windings are adjusted to maintain the end sections at a specified temperature, *i.e.*, the IPTS [15], as determined by two thermocouples positioned near the inner surface of these sections. The center winding is used to control the temperatures of the gold during gold-point calibrations.

Another high-temperature blackbody, used at NBS for accurate spectral calibrations of monochromators and spectrographs, has been used up to temperatures of 3273°K. It consists of a graphite cylindrical tube resistively heated in an argon atmosphere and surrounded by a number of graphite radiation shields. The tube is about 200 mm

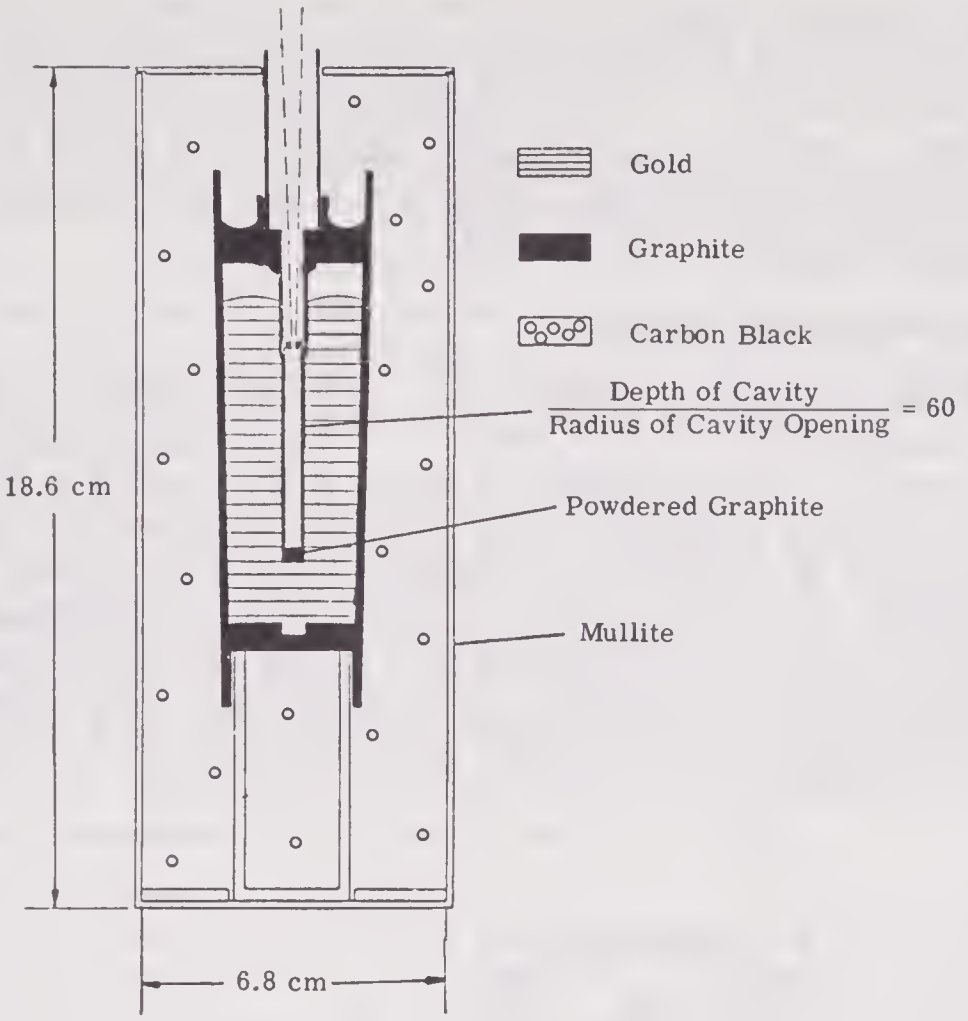


FIG. 3-4. NBS vertical blackbody.

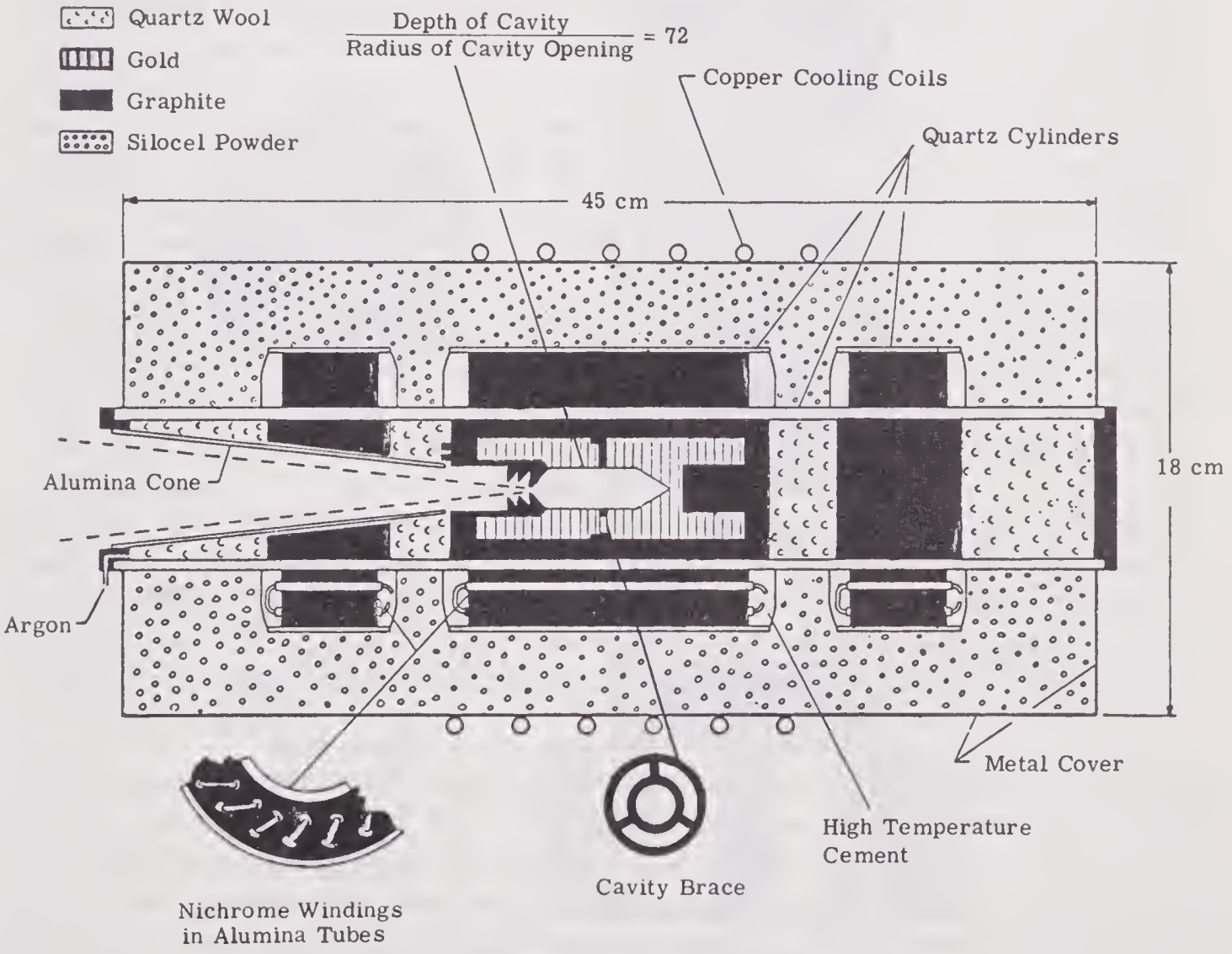


FIG. 3-5. NBS blackbody and furnace.

long, has a wall thickness of about 3 mm, and an inside diameter of about 9 mm. A small hole in the center of the tube and correspondingly larger holes in the shields permit the radiation to exit. A current of about 800 amp is required to reach a temperature of 3073°K. The tube can be used for about 50 hours at this temperature and the radiance can be stabilized for several hours to better than 1 percent by automatically controlling the current in the tube [14].

3.3. Laboratory and Field Sources

3.3.1. A High-Temperature Source [16]. "The high temperature element of the source is a graphite tube, with a small slit in one wall, heated directly with an alternating current. The arrangement is shown in cross section in Fig. 3-6. The graphite tube is rigidly supported at the upper end, and electrical connection is provided by means of a standard compression tube fitting modified by slightly enlarging its bore. At the other end, electrical connection and linear motion to relieve thermal stress are accomplished by a piston and bellows arrangement. The piston itself acts as the nut for another modified tube fitting; electrical insulation between the piston and the cylinder wall is produced by two O-rings, which

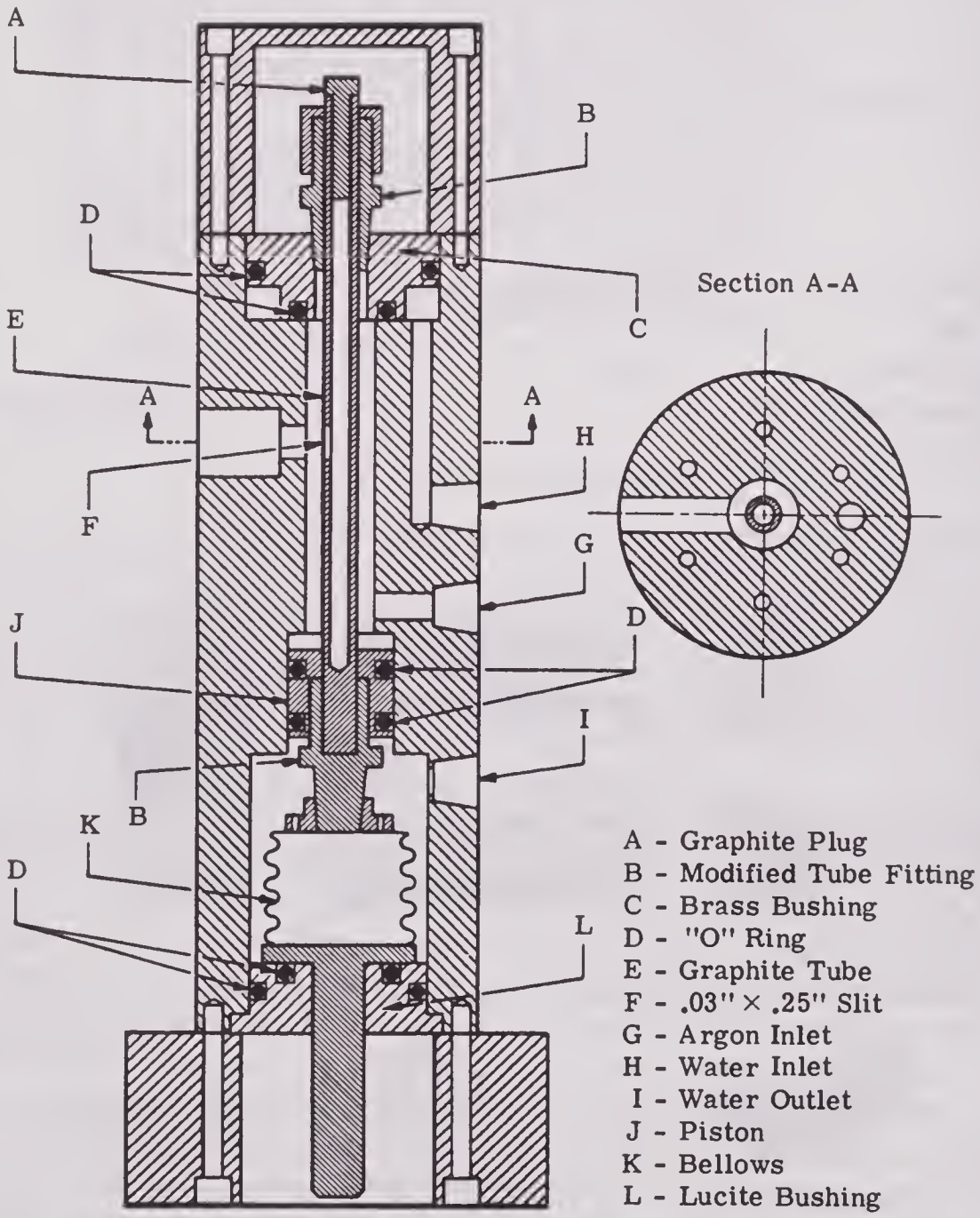


FIG. 3-6. Blackbody of Simmons, DeBell, and Anderson.

further serve as a fluid barrier. The bellows is mounted on a bushing of insulating material; electrical connection is made both to the central screw and the body wall. Both electrodes are cooled by water flow.

"The heated element was made from National Carbon Company type AGR graphite as a tube of 0.250-in. O.D. with a 0.050-in. wall; a slit of 0.25 × 0.03 in. was milled through one wall at approximately halfway between the points of first thermal or electrical contact. Erosion and oxidation of the hot graphite is reduced to a negligible rate by a flush and a slight flow of dry argon, which escapes through the body opening in front of the slit. The ends of the graphite tube were copper-plated to provide better contact.

"The power supply consists of two 20-amp toroidal autotransformers, one of which was converted into a conventional step-down transformer by removing the slider and adding a secondary winding of a few turns. Connection between this unit and the source body was made by means of a standard type of single-conductor insulated welding cable enclosed in 3/4-in. brass pipe which was connected to the base of the body. By this means, concentricity of the high current part of the electrical system is maintained; magnetic fields which would create spurious signals in nearby measurement circuits are thereby minimized.

"The current through the graphite and hence its temperature is controlled by manual adjustment of the unmodified autotransformer, the output of which is connected to the primary of the modified unit. The secondary current of the latter, *i.e.,* the current through the graphite tube, is measured using a 0-5 amp ac ammeter with a 40/1 current transformer. The time constant of the graphite tube is of the order of 10 sec so that no 60-cycle ac ripple is detectable in the radiant output. Constancy in source temperature is maintained by close control of line voltage with an electronic regulator, and by careful regulation of the argon flush.

"Calibrations were made with an optical pyrometer which in turn was checked against an NBS-certified tungsten ribbon lamp. The blackness of the source was not ascertained by experiment. However, an emissivity of 0.975 was computed assuming a uniform wall temperature, diffuse reflection from the inner wall, and an average surface emissivity of 0.85 for the graphite. (Source emissivities closer to unity can of course be obtained by reducing the ratio of the slit width to the internal tube diameter.)

"The highest temperature of operation is at present limited by the power supply to about 2400°K at which approximately 2 kw are being dissipated. During repetitive operation over periods of a few hours, the calibration has been found to be very stable following the first few heating and cooling cycles, during which a slight amount of further graphitization is presumed to have occurred. On occasion, over longer periods of many hours, appreciable drifts in the calibration have been observed. These shifts were found to be due to a gradual ablation of the graphite at a rate primarily related to the temperature of the graphite and the dryness of the argon. In the intermittent service for which the source was designed, however, the calibrations of the several units in use at Rocketdyne exhibit only very slight changes, which are monitored by occasional recalibrations."

3.3.2. Low-Temperature Large-Area Sources. Large-area thermal radiation sources used for comparison with field sources have been constructed [17]. One is an ambient-temperature conical field source, and another is a temperature-controllable conical field source. The ambient-temperature source consists of a simple sheet-metal cone coated internally with Sicon-black enamel. This source quickly assumes the ambient air temperature and is not temperature controlled. A thermometer

mounted so as to insure good thermal contact does the monitoring. The temperature distribution on the inner surface of the cone under normal conditions was determined with a surface pyrometer to be uniform within 0.5°C or less.

The temperature-controllable source consists of a similar cone with a water jacket. Provision is made for circulation and heating of the water. The impeller mounted near the apex and a stirrer blade are arranged such that a temperature uniformity of 1°K is maintained from the rear to the front of the cone as well as from top to bottom of the front portion of the source at an average temperature of 323° K.

Although no accurate experimental determinations have been made of the emissivity of these conical field sources, calculations have been made using the methods of DeVos [1] as applied to a conical source by Edwards [2]. The resulting value is 0.99.

Other field sources have been constructed commercially for use in comparing the radiance of large area targets. One in particular that has been used in the past is a 1 ft^2 flat metal surface covered with Zapon paint and temperature controlled, by means of a winding on the back, to a maximum temperature of about 250° C. The use of this "blackbody," manufactured by Barnes, and of most flat ones, is restricted because of the nonnegligible gradients of temperatures across the surface.

3.3.3. Nernst Glower. The Nernst glower is a hollow rod or filament about 3 cm long and 1 mm in diameter made from ZrO_2 (zirconia) and Y_2O_3 (yttria) mixed with CeO_2 (ceria) or ThO_2 (thoria). It is heated by an electric current from an approximately 180-w 50-v power supply. When used as a spectrometry source, the Nernst glower is intended to match entry slits about 12 mm high. The Nernst glower is generally operated at temperatures between 1500° and 2000°K, and has an expected life of about 200 to more than 1000 hours. Platinum wires are used as electrodes; therefore water cooling of end connections is not required. The connections are critical; construction details are given in *Rev. Sci. Inst.,* Vol. **11**, p. 429 (1940) [18] and *Phil. Mag.,* Vol. No. **50**, vi, p. 263 (1925) [19]. Because of its negative temperature coefficient of resistance, the Nernst glower must be heated to start and will continue to heat up until it burns out unless it is connected in series with a resistance to limit the current (ballasted).

The Nernst glower can be used for wavelengths up to about 30 μ; beyond 15 μ its emissivity decreases and its performance is comparable to that of the Globar or the Welsbach mantle. Low mechanical strength, small size, and susceptibility to temperature variations caused by air currents are disadvantages of the Nernst glower [20, 21].

3.3.4. Globar. The Globar is a rod of bonded silicon carbide. In its most common form it is about 5 cm long and 5 mm in diameter. It is heated electrically from a power supply of approximately 180 w at 50 v, and has a lifetime of about 250 hr when operated at 1500°K in air. It can be operated at temperatures up to 2200°K by a technique described by Strong [21]. Silver electrodes are used, and the ends of the Globar are coated with silver. A water-cooled jacket is used to prevent overheating of the end connections.

The Globar closely approximates a graybody, its emissivity being about 80% from 4 to 15 μ as shown in Fig. 3-7. Globar rods are available in large sizes for use as extended sources and for heating elements in furnaces and ovens operating up to about 1600°K [20].

3.3.5. Welsbach Mantle. The Welsbach mantle is a gauze similar to that used in gasoline lamps and lanterns. The gauze is impregnated with ThO_2 to which 0.75 to 2.5 percent CeO_2 is added. The gauze, or mantle, is heated either by burning gas

or an electric current (or both). The emissivity of the Welsbach mantle is relatively low up to about 6 μ, but between 10 and 100 μ it approximates a blackbody and is comparable to most other sources. At wavelengths exceeding 150 μ the emissivity falls off to the point where it is no longer usable. The Welsbach mantle can be operated at temperatures up to about 2400°K [20,21].

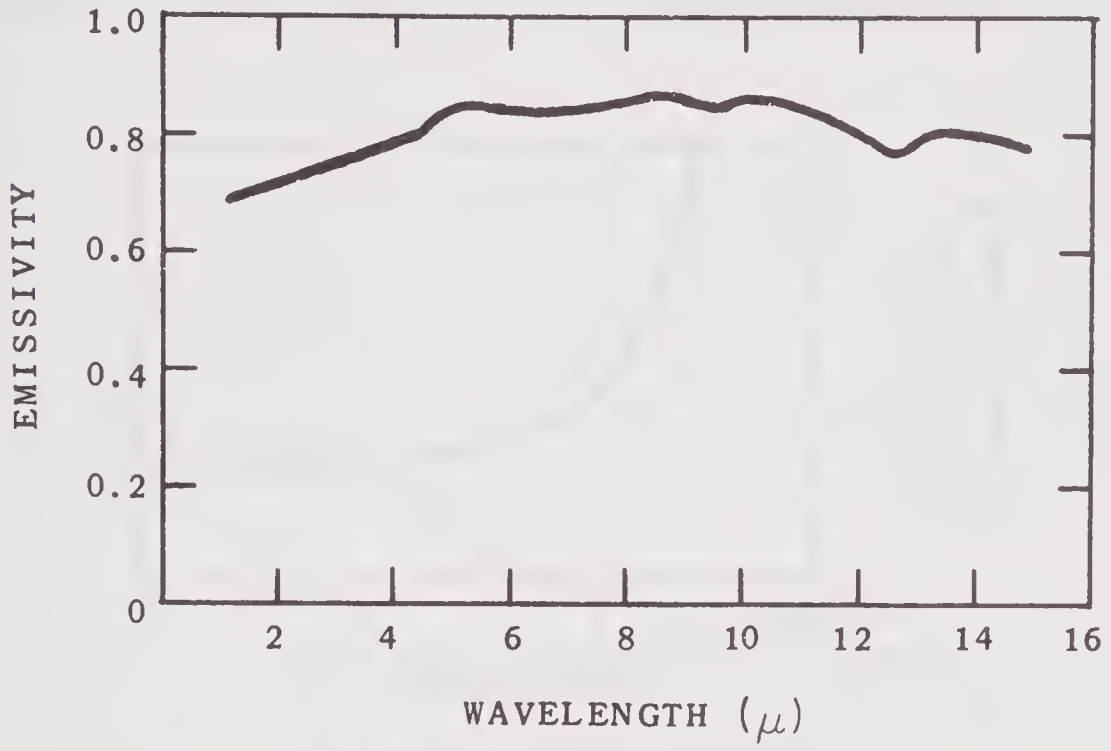


FIG. 3-7. Spectral output of Globar [22].

3.3.6. Carbon Arc. High source brightness is attained with the carbon arc, of which there are two main types: low intensity and high intensity. In the low-intensity arc, the radiation comes mainly from solid incandescent carbon in the shallow crater created on the face of the positive electrode. The temperature in this region is approximately 3900°K. In high-intensity arcs, a deeply formed crater appears in the specially cored positive electrode. This crater is characterized by a higher temperature and a higher current density than that in the low-intensity arc. Emission from vaporized core materials adds a dense visible line spectrum to the continuous spectrum emitted by the crater. The color temperature of the high-intensity arc can be as low as about 5000°K to as high as about 9000°K depending upon core material and current.

The electrode size, and therefore the current rating of a low-intensity arc, is optional since all low-intensity arcs have essentially the same positive crater temperature. The larger the electrode size, the larger the crater and the greater the ease of illuminating a given area, such as a spectrometer entrance slit. Typical carbon electrodes are 12 mm for the positive and 8 mm for the negative. A ballast resistor provides current stability with the negative coefficient of resistance of carbon.

The carbon arc is an excellent source in the wavelength range of 10 to 100 μ; beyond 100 μ a high-pressure mercury arc is better. Best source stability (better than $\pm 3\%$) is achieved by masking all but the crater; this is the most uniform part of the source. Details of operation of these sources are given in [23] and [24].

3.3.7. Tungsten Filaments. The tungsten-filament lamp, which operates at about 2800°K, is a source of high brightness. Because its glass envelope is not transparent

beyond about 3 μ, the tungsten-filament lamp, as normally used, is limited to about this wavelength. If, however, a tungsten filament is mounted in an enclosure behind a suitable infrared window, its useful range is limited only by the long-wavelength cutoff of the window. The signal from a tungsten filament mounted behind an alkali halide window was found [25] to be about one-half that of a carbon arc at all wavelengths from 2 to 14 μ. The ratio of the tungsten lamp emittance to that of a Globar is shown in Fig. 3-8 [25].

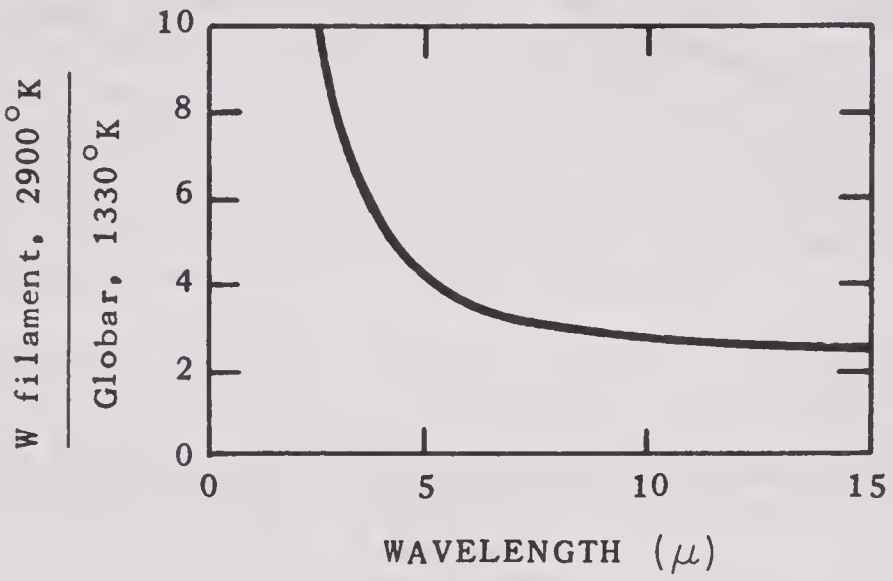


FIG. 3-8. Ratio of the output of a tungsten lamp at 2900°K to that of a Globar at 1330°K (replotted from [25]).

Another type of tungsten lamp called a strip- or ribbon-filament lamp is often used by NBS for spectral radiance calibrations of monochromators or spectrographs and as a comparison standard.

The G.E. Type 30A/6V/T24 tungsten strip lamp has a glass envelope about 75 mm in diameter and 300 mm long and a tungsten ribbon filament about 0.075 mm thick, 3 mm wide, and 50 mm long. The envelope contains argon at a pressure of about 0.3 atm at room temperature. This lamp requires a current of about 14 amp for a brightness temperature of 1273°K and about 45 amp for 2573°K. The change of brightness temperature with current varies from about 30 to 100°/amp from 1073° to 2573°K. Direct current is usually used so that standard potentiometric methods can be employed.

Tungsten lamps are commercially available from the General Electric Co. in the United States, the General Electric Co., Ltd., in England, and Phillips Lamp Works in the Netherlands.

3.3.8. Mercury Arcs. The chief laboratory source of far-infrared radiation is a commercially available high-pressure mercury arc operating in a fused-silica envelope. This source can be used for wavelengths from about 50 to 1400 μ. The high pressure (from 1 to 100 atm) of the arc broadens the discrete line spectrum into a continuous series of broad peaks. As the vapor pressure is increased, more of the emitted radiation occurs at longer wavelengths [21].

Figure 3-9 shows the ratio of the intensity of a quartz-mercury lamp compared to a Globar source. It can be seen that the energy of the mercury lamp is about six times greater at 200 μ and about three times greater at 100 μ. Between 20 and 50 cm^{-1} very little energy is emitted by the Globar. The quartz-mercury lamp used to obtain the curve was designated HPK 125 W and was made by Phillips' Lamp Works in the Netherlands. It had a working pressure of 3 atm. The lamp was operated at 135 v dc and a current of 0.98 amp. The Globar source was the usual type for commercial infrared spectrometers and was operated at 1200° K [26]. Other types of discharge lamps, notably H, Hg, and Xe lamps, can be used also as infrared sources.

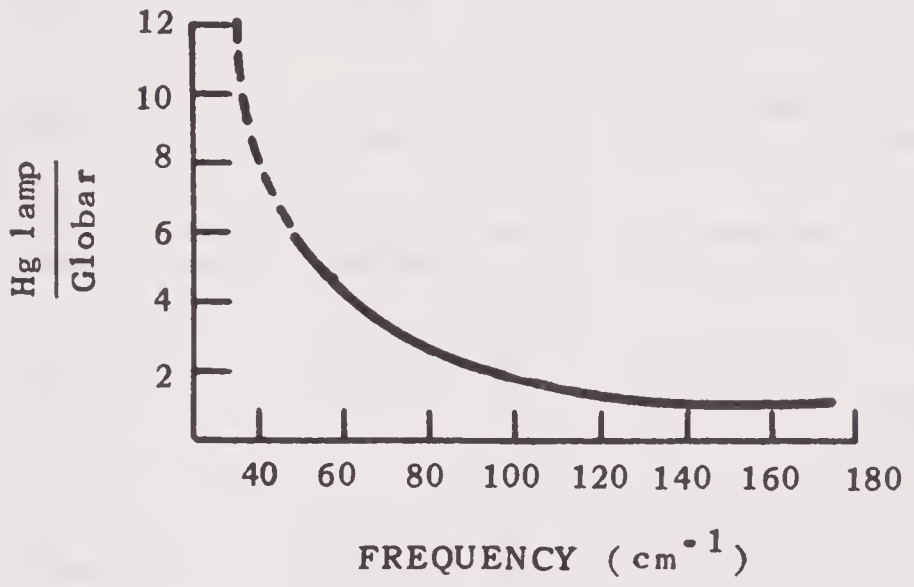


FIG. 3-9. Ratio of quartz-mercury lamp output to that of a Globar (replotted from [26]).

3.3.9. Zirconium Point Sources. An interesting type of arc, useful when a very small source of light is needed, is the so-called concentrated-arc lamp. The cathode consists of a small metal tube packed with zirconium oxide and the anode consists of a metal plate containing a hole slightly larger than the end of the cathode. Tungsten, tantalum, or molybdenum, because of their high melting points, are used for the metal parts. These are sealed in a glass bulb which is filled with an inert gas like argon to a pressure of nearly 1 atm. The arc runs between the (fused) surface of the zirconium oxide and the surrounding anode. The tip of the cathode is heated by ion bombardment to 2700°C or higher, giving it a surface brightness almost equal to that of the carbon arc. The light is observed through the hole in the anode. Lamps of this type can be made in which the source diameter is as small as 0.007 cm. These lamps are listed on pages 336 and 337 of Cenco Catalog J200. (They are probably also available elsewhere.) They come in powers rated from 2 to 300 w. These lamps must be ballasted; the Sylvania lamps come with operating instructions.

3.4. Commercial Cavity-Type Sources

Data on cavity-type blackbody sources currently available from various manufacturers are given in Table 3-3. Some can be obtained with aperture plates and other fittings.

TABLE 3-3. CURRENTLY AVAILABLE CAVITY-TYPE BLACKBODY SOURCES

Mfr.; Model No.*	*Type of Radiation Source*	*Temperature*		*Field of View* (degrees)	*Aperture Diameter*	*Emissivity*	*Warm-Up Time* (min); *Temp.* (°C)	*Max. Input Power* (w)	*Temp. Measuring Element*
		Range (°C)	*Accuracy* (°C)						
Barnes 11-100	14° conical cavity	60-230	±1[a]	20	0.50 in. 12.7 mm	0.99 ±1%	30 min 230°C	25	Th
Barnes 11-101	14° conical cavity	0-230	±1[a]	20	0.625 in. 16.9 mm	0.99 ±1%	30 230	100	Pt
Barnes 11-110	28° conical cavity	200-600	±3[a]	20	0.50 in. 12.7 mm	0.99 ±1%	120 600	400	Pt
Barnes 11-120	14° conical cavity	200-600	±3[a]	20	0.015 in. 0.38 mm	0.95+	2 200	55	Pt
Barnes 11-121	14° conical cavity	200-600	±3	20	0.04 in. 1 mm	0.95+	2 200	55	Pt
Barnes 11-131	14° conical cavity	200-1000	±5[a]	20	0.40 in. 10.2 mm	0.99 ±1%	60 1000	800	Pt
Perkin-Elmer PE 521-4 (Source) PE 521-5 (Controller)	15° conical cavity, blackened silver cone	50-600	±1[b]	20	0.50 in. 12.7 mm		20 600	160	Pt
Perkin-Elmer PE 521-6 (Source) PE 521-5R (Controller)	15° conical cavity, blackened silver cone	200-600	±2[c]	20	0.0087 in. 0.2 mm 0.015 in. 0.38 mm 0.040 in. 1.02 mm		<3 600	55	Pt
Perkin-Elmer PE 521-1 (Source) PE 521-5 (Controller)	15° conical cavity, blackened silver cone	200-600	±1.5[b]	20	0.040 in. 1.02 mm 0.015 in. 0.38 mm		<3 600	55	Pt
ITT	15° blackened conical cavity	200-600	±2		0.33 in. 8.4 mm			250	Pt
ITT	Blackened conical cavity	40-300	±1[d]		0.375 in. 9.5 mm				TC
IR Ind. IRI 403 (Source) IRI 101 (Controller)	20° blackened stainless steel cone	50-710	±1	14	0.0200 in. 5.08 mm 0.100 in. 2.54 mm 0.050 in. 1.27 mm 0.025 in. 0.64 mm 0.0125 in. 0.32 mm	0.99 ±0.01%	30	250	Pt

TABLE 3-3. CURRENTLY AVAILABLE CAVITY-TYPE BLACKBODY SOURCES (*Continued*)

Mfr.; Model No.*	*Type of Radiation Source*	*Temperature*		*Field of View* (degrees)	*Aperture Diameter*	*Emissivity*	*Warm-Up Time* (min); *Temp.* (°C)	*Max. Input Power* (w)	*Temp. Measuring Element*
		Range (°C)	*Accuracy* (°C)						
IR Ind. IRI 404 (Source) IRI 101 (Controller)	20° blackened stainless steel cone	50-1000	±1	14	0.200 in. 5.08 mm 0.100 in. 2.54 mm 0.050 in. 1.27 mm 0.025 in. 0.64 mm 0.0125 in. 0.32 mm	0.99 ±0.01%	45	525	Pt
IR Ind. IRI 405 (Source) IRI 101 (Controller)	20° blackened stainless steel cone	50-710	±1	30	0.600 in. 15.3 mm 0.400 in. 10.2 mm 0.200 in. 5.05 mm 0.100 in. 2.54 mm 0.050 in. 1.27 mm 0.025 in. 0.64 mm 0.0125 in. 0.32 mm	0.99 ±0.01%	60	525	Pt
IR Ind. IRI 406 (Source) IRI 102 (Controller)	Blackened conical cavity	200-600	±1	30	0.600 in. 15.3 mm 0.400 in. 10.2 mm 0.200 in. 5.05 mm 0.100 in. 2.54 mm 0.050 in. 1.27 mm 0.025 in. 0.64 mm 0.0125 in. 0.32 mm	0.99 ±0.01%	60	525	Pt
IR Ind. IRI 407 (Source) IRI 102 (Controller)	Blackened conical cavity	200-600	±1	10	0.039 in. 0.99 mm	0.99 ±0.01%	5	15	Pt

TABLE 3-3. CURRENTLY AVAILABLE CAVITY-TYPE BLACKBODY SOURCES (*Continued*)

Mfr.; Model No.*	*Type of Radiation Source*	*Temperature*		*Field of View* (degrees)	*Aperture Diameter*	*Emissivity*	*Warm-Up Time* (min); *Temp.* (°C)	*Max. Input Power* (w)	*Temp. Measuring Element*
		Range (°C)	*Accuracy* (°C)						
IR Ind. IRI 408 (Source) IRI 102 (Controller)	Blackened conical cavity	200-600	±1	90	0.100 in. 2.54 mm 0.050 in. 1.27 mm 0.025 in. 0.64 mm 0.0125 in. 0.32 mm	0.99 ±0.01%	5	60	Pt
IR Ind. IRI 411 (Source) IRI 101 (Controller)	Blackened conical cavity	710-1700	±1	14		0.99 ±0.01%	90	550	Pt
IRI Ind. IRI 412 (Source) IRI 106 (Controller)	Blackened conical cavity	50-900	±1	90	4.5×4.5 in. 114.3× 114.3 mm	0.90– 0.97	60	800	Pt
IRI Ind. IRI 417 (Source) IRI 103 (Controller)	Blackened conical cavity	50-1000	±1	18	2.0 in. 50.8 mm 1.5 in. 38.1 mm 1.0 in. 25.4 mm	0.99 ±0.01%	90	1100	Pt
IR Ind. IRI 420 (Source) IRI 101 (Controller)	Blackened conical cavity	200-1200	±1	14	0.200 in. 5.08 mm 0.100 in. 2.54 mm 0.050 in. 1.27 mm 0.025 in. 0.64 mm 0.0125 in. 0.32 mm	0.99 ±0.01%	45	380	Pt
IR Ind. IRI 424 (Source) IRI 101 (Controller)	Blackened conical cavity	710-1700	±1	14	0.050 in. 1.27 mm 0.025 in. 0.64 mm 0.0125 in. 0.32 mm	0.99 ±0.01%	90	550	Pt
IRI Ind. IRI 427 (Source) IRI 102 (Controller)	Blackened conical cavity	200-600	±1	14	0.080 in. 2.03 mm	0.99 ±0.01%	3	15	Pt
Radiation Elec.	V-grooved metal block	60-520	±2		2.125 in. 54.0 mm	0.99 ±0.01%	30 1000° F	225	TC (Fe-constantan)

TABLE 3-3. CURRENTLY AVAILABLE CAVITY-TYPE BLACKBODY SOURCES (*Continued*)

Mfr.; Model No.*	*Type of Radiation Source*	*Temperature*		*Field of View* (degrees)	*Aperture Diameter*	*Emissivity*	*Warm-Up Time* (min); *Temp.* (°C)	*Max. Input Power* (w)	*Temp. Measuring Element*
		Range (°C)	*Accuracy* (°C)						
Elec. Communications	Cylindrical blackbody cavity	Ambient to 1000 continuously adjustable	±1		1.125 in. 28.6 mm	0.995	70 1000	700 max.	TC in air
Eppley Labs.	Blackened stainless steel cavity	600-1100	±1		0.75 in. 19.1 mm	0.97	,120 approx. 1000	1500	TC (Pt-Pt 10% Rh)
Williamson	Cylindrical blackbody cavity	Ambient to 65	±1		0.75 in. 19.1 mm		60 65	<10	Mercury-in-glass thermometer

*Barnes = Barnes Engineering Company.
Perkin-Elmer = Perkin-Elmer Corporation.
ITT = International Telephone and Telegraph Corporation.
IR Ind. = Infrared Industries.
Radiation Elec. = Radiation Electronics Company.
Elec. Communications = Electronics Communications, Inc.
Eppley Labs. = Eppley Laboratories, Inc.
Williamson = Williamson Development Company.

[a]Maximum error due to combined shift of calibration, ambient temperature (0° to 40°C) and line voltage (105 to 125 v).
[b]Temperature variations from 16° to 38°C and line-voltage variation from 105 to 125 v, 60 cycles ac.
[c]Temperature variations from 20° to 40°C and line-voltage variation from 105 to 125 v, 60 cycles ac.
[d]At 300°C with smaller variation at lower temperature.

References

1. J. C. DeVos, *Physica,* **20**, 669 (1954).
2. David F. Edwards, "The Emissivity of a Conical Blackbody," 2144-105-T, The University of Michigan Engineering Research Institute (now Institute of Science and Technology, The University of Michigan), Ann Arbor (1956).
3. Andre Gouffé, *Rev. Optique,* **24**, 1 (1945).
4. C. S. Williams, *J. Opt. Soc. Am.,* **51**, 564 (1961).
5. E. M. Sparrow, L. U. Albers, and E. R. G. Eckert, *J. Heat Transfer*, 73 (1962).
6. G. J. Zissis, "Precision Radiometry-Theory," *Special Topics in Infrared Technology,* Engineering Summer Conferences, The University of Michigan, Ann Arbor (1963).
7. E. W. Truenfels, *J. Opt. Soc. Am.,* **53**, 1162 (1963).
8. E. M. Sparrow, L. V. Albers, and E. R. G. Eckert, *J. Heat Transfer Trans. ASME* (1962).
9. E. M. Sparrow and L. V. Albers, *J. Heat Transfer Trans. ASME,* Ser C, **82**, 233 (1960).
10. *NBS Bulletin No. 227,* **11**, 87 (1914).
11. *NBS J. Research No. 578,* **11**, 79 (1933).
12. *NBS J. Research,* **53**, 211 (1954).
13. W. N. Harrison, Joseph C. Richmond, Earle K. Plyler, Ralph Stair, and Harold K. Skramstad, "Standardization of Thermal Emittance Measurements," WADC TR 59-510, National Bureau of Standards (1960) AD 238 918.
14. H. J. Kostowski and R. D. Lee, "Theory and Methods of Optical Pyrometry," *NBS Monograph 41,* U. S. Department of Commerce, National Bureau of Standards.
15. H. F. Stimson, *J. Research Natl. Bur. Standards,* **42**, 209 (1949); *J. Research Natl. Bur. Standards,* **65A**, 139 (1961).
16. F. Simmons, A. G. DeBell, and Q. S. Anderson, *Rev. Sci. Instr.,* **32**, 1265 (1961).
17. A. LaRocca and G. Zissis, *Rev. Sci. Instr.,* **30**, 200 (1959).
18. *Rev. Sci. Inst.,* **11**, 429 (1940).
19. *Phil. Mag.,* **50**, 263 (1925).
20. P. W. Kruse, L. D. McGlauchlin, and R. B. McQuistan, *Elements of Infrared Technology-Generation, Transmission, and Detection,* Wiley, New York (1962).

21. J. Strong, *Procedures in Experimental Physics,* Prentice-Hall, Inc., New York (1939).
22. R. A. Smith, F. E. Jones, and R. P. Chasmar, *The Detection and Measurement of Infra-Red Radiation,* Oxford University Press, Oxford (1957), originally from S. Silverman, *J. Opt. Soc. Am.,* **38**, 989 (1948).
23. *Symposium on Searchlights,* The Illuminating Engineering Society, 32 Victoria St., London SW1, printed by Argus Press, Ltd., Temple Ave. and Tudor St., London EC4, England.
24. W. Finkelnberg, "The High Current Carbon Arc and Its Mechanism," *J. Appl. Phys.,* **20,** 468 (1949).
25. J. H. Taylor, C. S. Rupert, and J. Strong, *J. Opt. Soc. Am.,* **41,** 626 (1951).
26. E. K. Plyler, D. J. C. Yates, and H. A. Gebbie, *J. Opt. Soc. Am.,* **52,** 859 (1962).

Chapter 4

TARGETS

C. E. Dunning

Aerojet-General Corporation

Fred E. Nicodemus

Sylvania Electronic Systems

CONTENTS

4.1. Introduction 58
4.2. Aerial Targets 58
4.2.1. Symbols, Definitions, and Equations 58
4.2.2. Jet Structure 60
4.2.3. Exhaust Composition 62
4.2.4. Heating During Exit 68
4.2.5. Heating During Reentry 69
4.3. Surface Targets 69
4.4. Radiometric Analysis and Discussion 69
4.4.1. Thermal Emission 70
4.4.2. Reflection 72
4.4.3. Factors Affecting the Temperatures of Passive Targets 72
4.5. Measured Values of Radiant Emissivity and Reflectance 74

4. TARGETS

4.1. Introduction

Although a great deal of military infrared technology does not differ greatly from infrared technology in general, and can be freely treated in an unclassified work such as this Handbook, security classification forbids such discussion in certain areas. The characteristics of military infrared targets is just such an area, so that this chapter is unavoidably incomplete. About all that can be said is that, within this severe limitation, an attempt has been made to include as much information as possible.

In the first part of this chapter are data concerning rocket engines, which may be helpful in attempts to estimate their characteristics as infrared targets. In the second part radiometric principles are discussed as they apply to the radiation from a wide variety of targets, particularly of surface targets as they appear in the output of an infrared scanner. Also included are data regarding the emissivities of a number of common objects and surfaces.

4.2. Aerial Targets (Rocket-Propulsion Systems)

Aerial targets may be military objectives; they may be aircraft or rockets which are a threat to security, or they may be spacecraft which must rendezvous or track and home on a planet. The power plants can be piston engines, jets, or rockets. The radiation comes from the exhausts of the engines, from frictional heating of the structure passing through the air, and from the air heated by very high speed objects like reentry bodies.

4.2.1. Symbols, Definitions, and Equations

A = nozzle cross-section area (in.^{2})
c = effective exhaust velocity (ft sec^{-1})
c_h = Stanton number (proportional to local skin-friction coefficient, highly dependent on state of flow, *i.e.*, laminar or turbulent; usually 1 to 2×10^{-3})
c_p = specific heat at constant pressure
d = nozzle diameter
F = thrust (lb)
H_A = specific enthalpy (total internal energy per unit mass) of ambient atmosphere
H_E = specific enthalpy of unmixed jet
H_j = specific enthalpy (total internal energy per unit mass) of mixed jet
I_{sp} = specific impulse (lb-sec lb^{-1}). Varies approximately as $T/\mathfrak{M}^{-1/2}$.
K_p = chemical-equilibrium constant
m = proportion of entrained atmosphere
M = Mach number, unless otherwise noted
$\mathfrak{M}$ = molecular weight
P = pressure (lb in.$^{-2}$)
R = specific gas constant (ft-lb lb^{-1} °R^{-1}), unless otherwise noted

T = absolute static temperature in °R unless otherwise noted
v = velocity (ft sec^{-1})
V_m = the differential velocity between the initial jet and the atmosphere
$\dot{W}$ = exhaust-gas weight flow rate (lb sec^{-1})
γ = ratio of specific heats (c_p/c_v); depends basically on the number of atoms in the molecule. For monatomic gases, $\gamma = 1.66$; for diatomic gases, $\gamma \approx 1.30$.
ρ = air density outside boundary layer
μ = air velocity outside boundary layer

Subscripts

a = ambient
c = combustion chamber
e = nozzle exit plane
t = nozzle throat
w = wall

THRUST AND MASS FLOW

$$F = \frac{\dot{W}_c}{g} = \frac{\dot{W}v_e}{g} + A_e(P_e - P_a) = I_{sp}\dot{W}$$

EXHAUST VELOCITY

$$v_t = (\gamma g R T_t)^{\frac{1}{2}}$$

$$v_e \approx c = I_{sp}g$$

$$M = v/(\gamma g R T)^{\frac{1}{2}}$$

$$v = \left[\frac{2\gamma g R}{\gamma - 1}(T_c - T)\right]^{\frac{1}{2}}$$

TEMPERATURE

$$T_t = \frac{2}{\gamma + 1} T_c$$

$$T = T_c\left[1 + \frac{\gamma - 1}{2} M^2\right]^{-1}$$

$$= T_c(P/P_c)^{(\gamma - 1)/\gamma}$$

These three equations, relating temperature of the jet exhaust to internal conditions of the motor, are not precisely fulfilled in practice because of factors such as conductive and radiative heat losses, changing chemical composition, and nonideal-gas behavior.* Boynton and Neu [1] found that actual exhaust temperatures were

*The energy budget of a typical rocket engine is as follows: 1% loss due to incomplete combustion, 2% heat loss to engine walls, 27-57% in unavailable jet thermal energy, the balance in useful propulsion energy and residual exhaust kinetic energy.

best predicted by assuming shifting chemical equilibrium down to a certain pressure and frozen composition thereafter; this so-called "freezing pressure" was determined by fitting to empirical data. Either process assumed alone would have incurred errors of about 200°K.

Mixing and Afterburning. The energy balance of an exhaust jet of frozen composition, mixing with the ambient atmosphere, has the following form [1]:

$$H_j = \frac{H_E + mH_A}{1+m} + \frac{m(v_j - v_m)^2}{2g(1+m)^2}$$

More heat may be added to the jet by further reactions with the entrained atmosphere (afterburning). Table 4-1 gives the calculated heats released by afterburning processes for several fuels [2].

Table 4-1. Calculated Heat Release During Afterburning

Propellant Combination	O/F (by wt)	P_c, P_e (psia)*	%H_2 (by wt)	%CO (by wt)	*Heat Release*† (Btu/lb propellant)
LO_2/JP-4	2.127	600/14.7	18.7	37.7	3117
N_2O_4/$\frac{1}{2}$UDMH, $\frac{1}{2}N_2H_4$	2.01	766/14	3.1	3.4	358
LO_2/NH_3	1.41	600/14.7	4.8	–	305
LO_2/C_2H_5OH	1.5	300/14.7	10.2	24.7	1916
LO_2/N_2H_4	0.66	600/14.7	22.0	–	1540
WFNA/aniline-alcohol	3.0	315/14.7	5.1	21.9	1234
RFNA/N_2H_4	1.159	300/14.7	14.4	–	914
LO_2/LH_2	5.556	600/14.7	29.4	–	2830
LF_2/LH_2	5.65	600/14.7	5.4	–	6420
LF_2/N_2H_4	2.4	600/14.7	0.3	–	20
Solid C_nH_{2n} Al, NH_4ClO_4	–	1000/14.7	41.4	29.8	3706

*Underline indicates pressure to which the given gas composition pertains.
†Considering complete combustion of H_2 and CO.

Stagnation. Ideally, stagnation temperatures of an isentropic exhaust stream would equal the initial or chamber temperature, T_c. Figure 4-1 illustrates representative stagnation temperatures [3].

4.2.2. Jet Structure. The region behind a rocket or jet engine is an extremely complicated chemical and thermodynamic entity. This jet structure has been analyzed in great detail because knowledge of how much of what kind of gas is where and at what temperature makes possible calculations of the radiation field. This region, also called the flow field, is illustrated in Fig. 4-2. The flow-field characteristics are functions of both the engine and nozzle characteristics as well as the atmosphere in which the motor is operating.

Thrust in Vacuum	171,000 lb
Thrust at Sea Level	150,000 lb
Nozzle Area Ratio	8:1
Nozzle Exit Diameter	4.3.14 in.
Propellants	LO_2-RP-1
Chamber Pressure	558 PSIA
Nozzle Exit Pressure	10.3 PSIA
Nozzle Exit Static Temp	3400°F
Nozzle Exit Stagnation Temp	6000°F

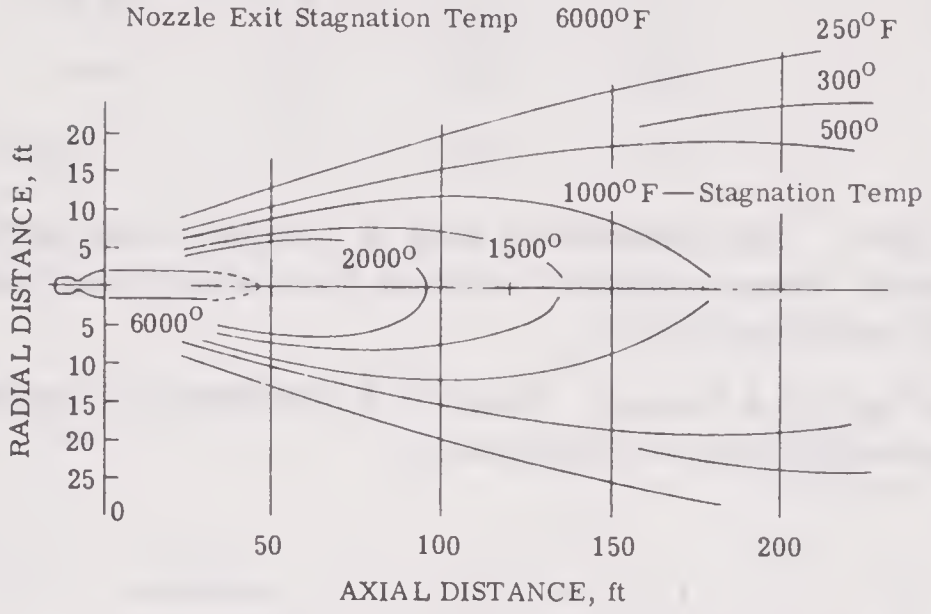


FIG. 4-1. Stagnation temperature (°F) distribution in exhaust at sea level.

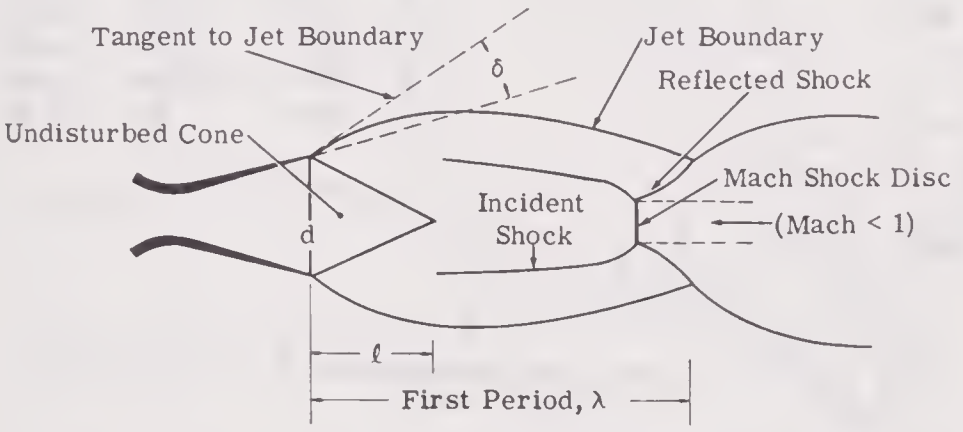


FIG. 4-2. Features of the flow field.

4.2.2.1. *Flow Field Within the Atmosphere.* Figure 4-2 actually illustrates the flow field for operation in the atmosphere for different nozzle expansion ratios and pictorially defines the shock, first period, and undisturbed cone.

Aerodynamic heating is discussed in Chapter 21.

DEPARTURE OF THE JET FROM THE NOZZLE. δ = the angle at the nozzle exit between the tangent to the nozzle and the tangent to the jet boundary.

Underexpanded nozzle	$P_e > P_a$	$\delta > 0$
Optimally expanded nozzle	$P_e + P_a$	$\delta > 0$
Overexpanded nozzle	$P_e < P_a$	$\delta < 0$

SHOCKS. For shock formation in supersonic jets, see Fig. 4-2. The theory of shock formation is covered in [5] and [6].

FIRST PERIOD. The first period of the jet is measured from the nozzle exit to the beginning of the reflected shock. See Table 4-2.

TABLE 4-2. LENGTH IN NOZZLE DIAMETERS [4]

Exit Mach No. ($\gamma = 1.4$)	P_e/P_a				
	0.5	*1.0*	*2.0*	*4.0*	*7.0*
1.5	0	1.2	4.0	5.0	8.0
2.5	0.8	2.4	4.0	5.6	8.8
3.5	1.6	3.4	5.6	8.4	–
4.5	2.1	4.4	7.2	–	–

UNDISTURBED CONE. The undisturbed cone is a region whose temperature, composition, and length remain relatively constant for a given nozzle configuration ($1 \approx 3/2d$) under all conditions of P_a [7].

4.2.2.2. *Flow Field in a Vacuum.* Figure 4-3 illustrates the expansion of an exhaust in a vacuum under several conditions [5].

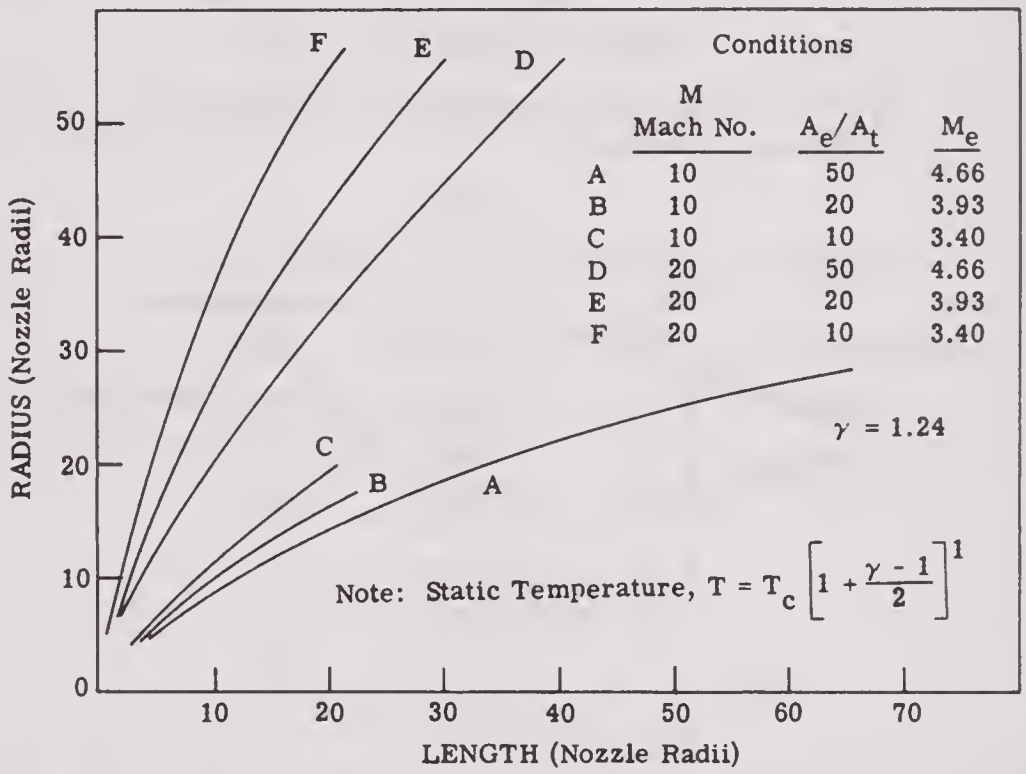


FIG. 4-3. Exhaust expansion into a vacuum; line of constant Mach number [6].

4.2.3. Exhaust Composition

4.2.3.1. *Gases.* Table 4-3 lists a number of liquid propellants and their combustion products. Lox/RP-1 exhaust composition *vs* mixture ratio is plotted in Fig. 4-4; Fig. 4-5 is a similar plot for $N_2O_4/0.5N_2H_4 + 0.5$UDMH.

CHEMICAL EQUILIBRIUM. The balance of reactants and products achieved in chemical equilibrium is expressed as

$$K_p = \frac{P_1 \cdot P_2 \cdot P_3 \cdots}{P_1 \cdot P_m \cdot P_n \cdots}$$

where the P terms in the numerator are the partial pressures of the individual exhaust gases and the P terms in the denominator are the partial pressures of the reactant gases.

Figure 4-6 [6] shows the equilibrium conditions as a function of temperature for a number of important reactions.

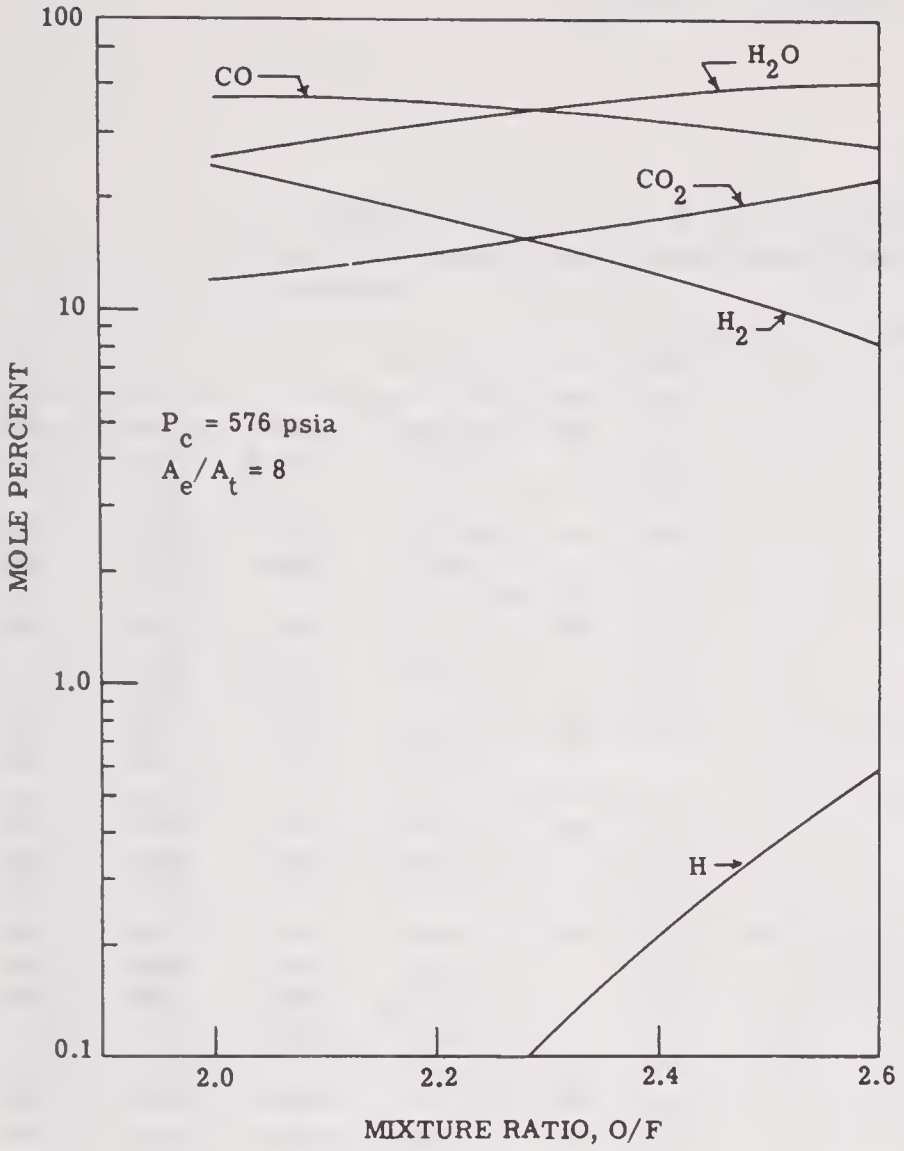


FIG. 4-4. Exhaust composition of liquid oxygen/RP-1 *vs* mixture ratio.

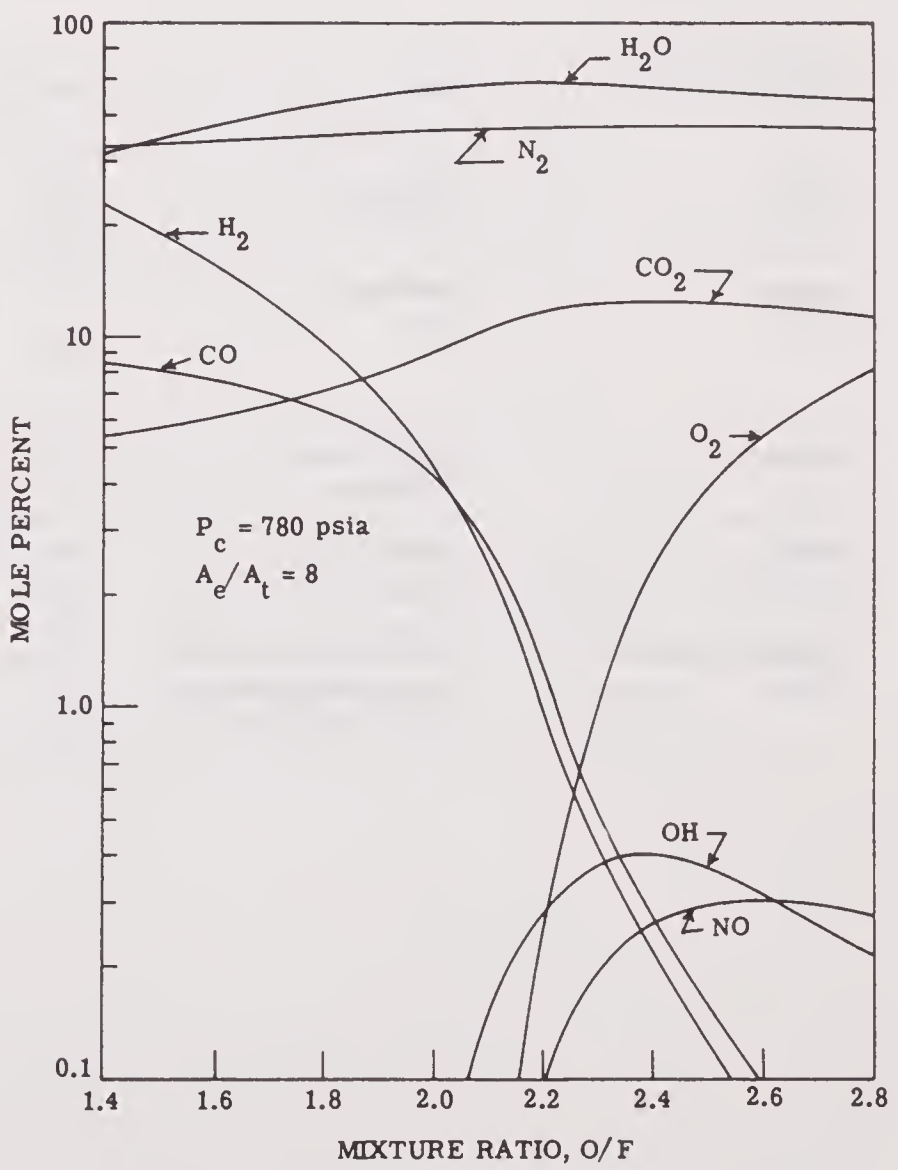


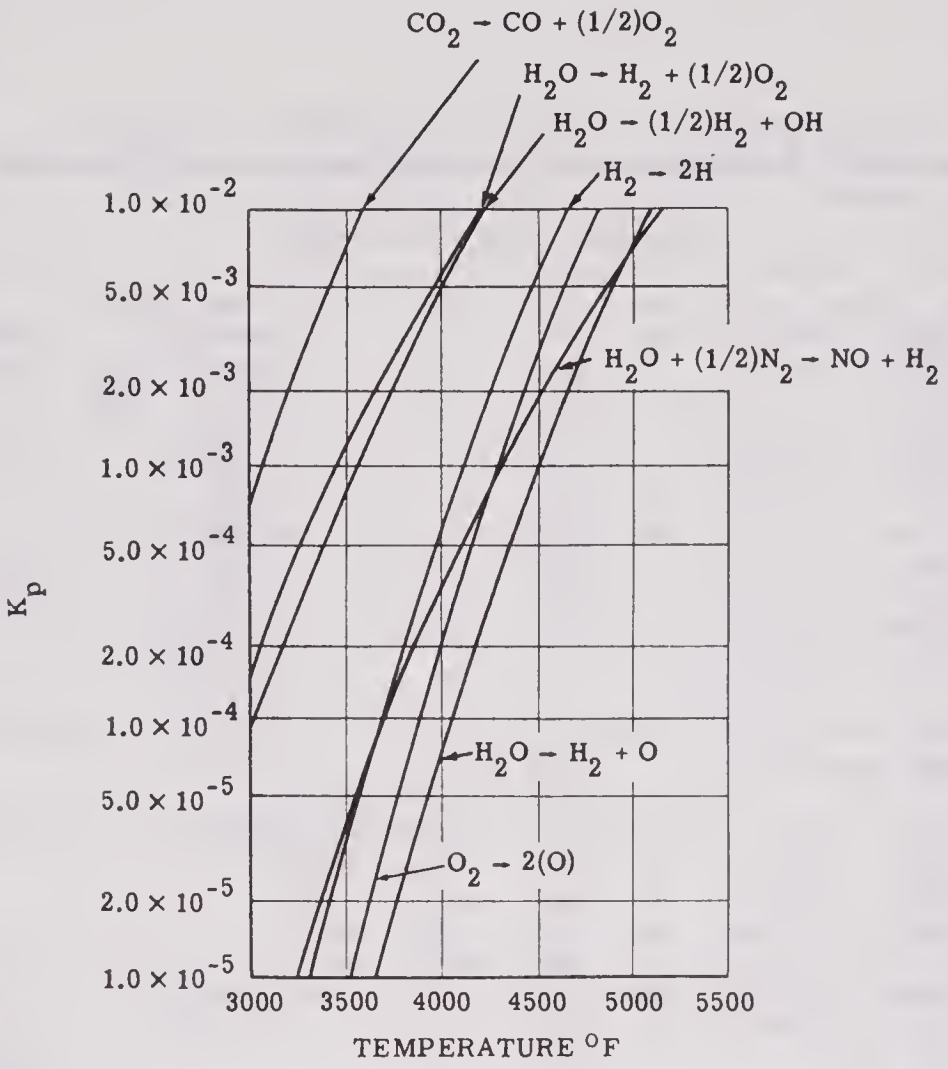
FIG. 4-5. Exhaust composition of $N_2O_4/0.5N_2H_4$ + 0.5UDMH *vs* mixture ratio.

TABLE 4-3. REPRESENTATIVE LIQUID PROPELLANTS

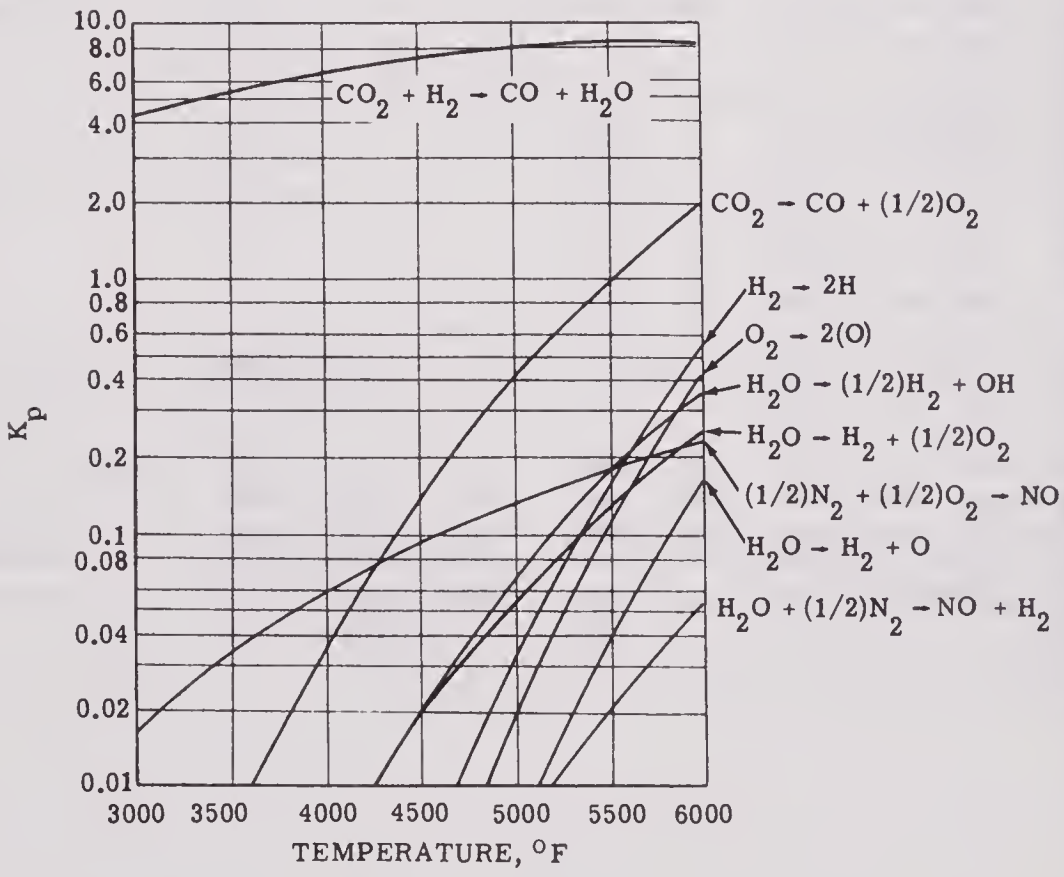
Oxidizer	*Fuel*	*Conditions*					
		O/F	P_c	P; A_e/A_t	T_c	T	I_{sp}
Chlorine Trifluoride (CLF_3)	Hydrazine (N_2H_4)	2.6	300	14.7;3.6	6597°A	4139°R	258
				0.3;53	6597°A	1588	330
Fluorine (F_2)	Ammonia (NH_3)	3.2	300	14.7;3.9	7760	5521	312
				0.6;37	7760	2829	403
Fluorine (F_2)	Hydrazine (N_2H_4)	2.2	300	14.7;3.9	4415	3179	316
				0.6;39	4415	1710	411
Fluorine (F_2)	Hydrogen (H_2)	6.0	100	14.7;1.8	5913	4480	303
				0.2;149	5913	815	484
98% Hydrogen Peroxide (H_2O_2)	Hydrazine (N_2H_4)	1.8	300	14.7;3.7	5152	3249	253
				0.6;38	5152	1718	325
Nitrogen Tetroxide (N_2O_4)	1/2 UDMH, 1/2 Hydrazine [$(CH_3)_2N_2H_2$, N_2H_4]	1.8	800	14.7;7.5	5946	3248	282
				0.8;66	5946	1830	344
		2.0	30	14.7;1.02	5389	5079	127
				0.3;76	5389	2143	335
Nitrogen Tetroxide (N_2O_4)	Pentaborane (B_5H_9)	3.2	300	14.7;4.1	6812	5298	259
				0.3;98	6812	3754	364
Oxygen (O_2)	RP-1 ($C_{10}H_{20}$)	2.2	800	14.7;7.6	6287	3474	291
				1.6;40	6287	2306	339
Oxygen (O_2)	Hydrazine (N_2H_4)	0.8	500	14.7;5.4	5915	3567	292
				1.0;38	5915	2068	354
Oxygen (O_2)	Hydrogen (H_2)	3.0	300	14.7;3.4	4837	2660	364
				0.3;51	4837	1033	462
		3.0	50	14.7;1.2	4734	3862	251
				0.05;51	4734	1036	461
Oxygen (O_2)	Ethyl Alcohol (C_2H_5OH)	1.5	300	14.7	5705	3170	242
IRFNA ($HNO_3 + NO_2 + H_2O + HF$)	UDMH ($(CH_3)_2N_2H_2$)	2.8	206	14.7;2.9	5400	3857	226
				.21;54		1075	293
Nitrogen Tetroxide (N_2O_4)	15% Nitric Oxide, 85% Mono Methyl Hydrazine [(NO, $N_2H_3(CH_3)$]	2.1	30	14.7;1.02	5454	5135	129
				0.01;252	5454	1487	359

TABLE 4-3. REPRESENTATIVE LIQUID PROPELLANTS *(Continued)*

Combustion Products (mole percent)											
H_2O	H_2	H	OH	O	NO	N_2	CO_2	CO	HF	HCl	*Other*
–	4.3	0.2	–	–	–	20.7	–	–	56.1	18.1	0.61 Cl
–	4.1	–	–	–	–	20.8	–	–	56.3	18.8	–
–	1.33	2.22	–	–	–	14.3	–	–	81.0	–	1.11 F
–	1.19	–	–	–	–	14.6	–	–	83.5	–	–
–	1.82	3.57	–	–	–	20.1	–	–	73.1	–	1.4 F
–	2.97	0.01	–	–	–	20.6	–	–	76.4	–	–
–	50.4	1.71	–	–	–	–	–	–	47.9	–	–
–	51.7	–	–	–	–	–	–	–	48.3	–	–
72.2	6.48	0.01	0.02	–	–	21.3	–	–	–	–	–
72.2	6.48	–	–	–	–	21.3	–	–	–	–	–
42.3	9.45	0.01	–	–	–	34.9	7.15	6.20	–	–	–
38.8	13.0	–	–	–	–	34.9	10.7	2.70	–	–	–
38.3	6.81	2.30	3.32	0.84	0.75	33.5	5.43	6.93	–	–	1.8 O_2
44.5	6.32	–	–	–	–	36.0	10.9	2.19	–	–	–
3.07	25.1	6.62	0.23	0.04	0.02	22.3	–	–	–	–	5.99 BO, 1.41 B_2O_2
1.01	33.3	1.15	0.01	–	–	23.7	–	–	–	–	28.3 HBO_2, 6.87 B_2O_3
											0.47 BO, 0.56 B_2O_2
											27.3 HBO_2, 12.5 B_2O_3
32.1	17.8	0.04	0.01	–	–	–	14.3	35.7	–	–	–
26.7	23.3	–	–	–	–	–	19.7	30.3	–	–	–
53.4	13.2	0.05	0.02	–	–	33.3	–	–	–	–	–
53.4	13.2	–	–	–	–	33.3	–	–	–	–	–
37.8	62.2	–	–	–	–	–	–	–	–	–	–
37.8	62.2	–	–	–	–	–	–	–	–	–	–
37.7	61.9	0.32	0.01	–	–	–	–	–	–	–	–
37.8	62.2	–	–	–	–	–	–	–	–	–	–
44.9	10.2	1.90	3.37	0.41	–	–	13.8	24.7	–	–	0.67 O_2
48.0	6.24	0.10	0.08	–	0.02	22.0	4.74	17.5	0.53	–	–
35.3	19.1	0.59	0.24	0.01	0.02	22.0	4.74	17.5	0.53	–	–
34.8	8.55	2.88	3.06	0.76	0.62	32.7	5.63	9.82	–	–	1.18 O_2
35.2	13.7	–	–	–	–	34.8	14.7	1.63	–	–	–



(a)



(b)

FIG. 4-6. Equilibrium constant K_p *vs* temperature [6].

Molecular Emission. Table 4-4 is a tabulation of the infrared emission bands of the principal exhaust gases.

Table 4-4. Major Infrared Emission Bands of Exhaust Gases [8]

Gas	*Approximate Center of Major Emission Bands* (μ)
H_2O	0.94, 1.14, 1.38, 1.88, 2.66, 2.74, 3.17, 6.27
CO_2	1.96, 2.01, 2.06, 2.69, 2.77, 4.26, 4.68, 4.78, 4.82 5.17, 15.0
CO	4.663, 2.345, 1.573
HCl	3.465, 1.764, 1.198
NO	5.30, 2.672
NO_2	4.50, 6.17, 15.4
N_2O	2.87, 3.90, 4.06, 4.54, 7.78, 8.57, 16.98
OH	1.00, 1.03, 1.08, 1.14, 1.21, 1.29, 1.38, 1.43, 1.50, 1.58, 1.67, 1.76, 1.87, 1.99, 2.15, 2.80, 2.94, 3.08, 3.25 3.43, 3.63, 3.87, 4.14, 4.47
SO_2	4.0, 4.34, 5.34, 7.35, 8.69

4.2.3.2. *Particles in Exhausts.* Boynton [9] measured the carbon particle size distribution in the exhaust of a small rocket engine. The distribution is shown in Fig. 4-7. It shows the apparent diameters in 22,000× micrographs.

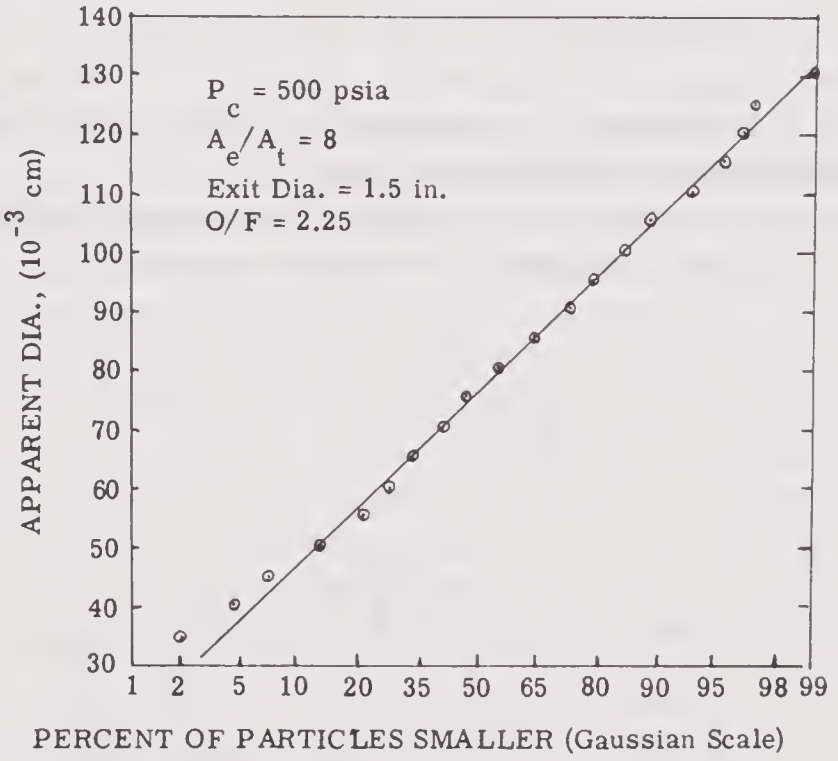


Fig. 4-7. Size distribution of carbon particles.

Kurtovich and Pinson [10] measured the particle-size distribution in the exhaust stream of a scale-model solid-propellant rocket. In this case, the particles were oxide products of the pulverized aluminum used in some high-energy solid fuels. The distributions at two points downstream of the exit are given in Fig. 4-8. In Fig. 4-9 the computed effect of particle size on the overall exhaust emissivity is given, for the distribution indicated by the solid line in Fig. 4-8.

In Fig. 4-8 the particle size distribution at two distances from the nozzle exit plane is given. The volumetric particle concentration was 1.14×10^{-5}. The solid curve

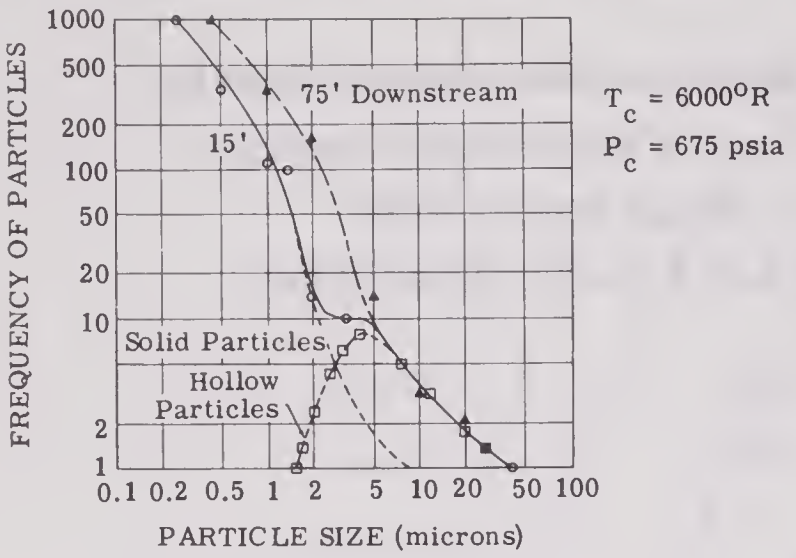


FIG. 4-8. Particle size distribution.

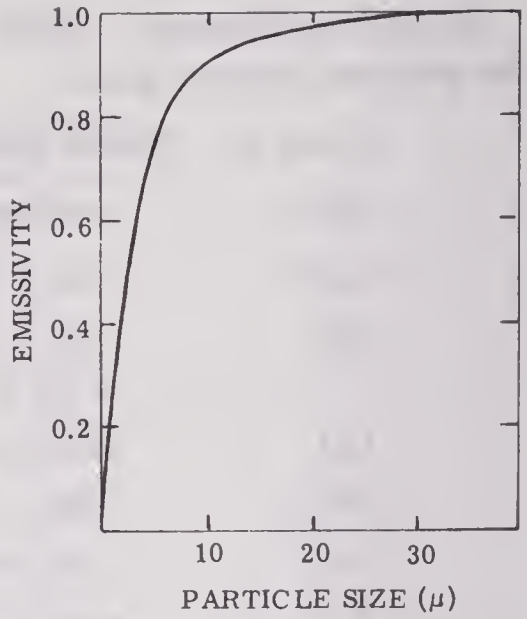


FIG. 4-9. Emissivity as a function of upper limit of particle size.

in Fig. 4-8 represents a distribution for which the emissivity of a particle cloud 1 ft thick has been calculated. This is shown in Fig. 4-9 as a function of an upper limit to particle size.

4.2.4. Heating During Exit. Heating during exit, as a result of supersonic velocities in the atmosphere, may be calculated, though certain data, particularly those regarding the boundary-layer flow, are not completely determined. The heat flow, q, into a surface is calculated as follows:

$$q = c_h \rho \mu c_p \left[T\left(1 + R\frac{\gamma - 1}{2}M^2\right) - T_w \right]$$

In this equation, T = boundary-layer temperature and R is the recovery factor (0.82 to 0.88 [11] depending mostly on the state of flow).

RADIATION PROCESSES. Aerodynamic heating is discussed in Chapter 21. Figure 4-10 is an example of a time-temperature curve of a nose fairing.

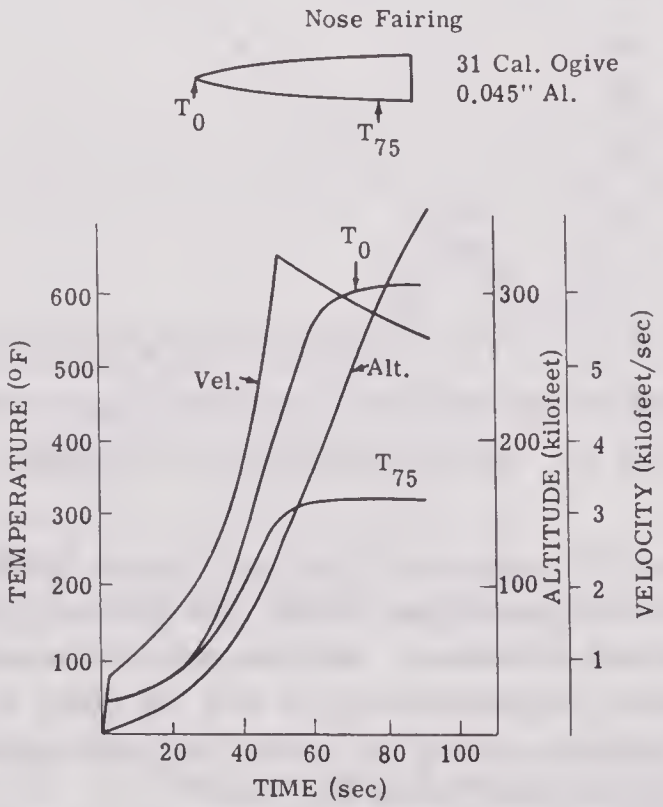


FIG. 4-10. Temperature of Aerobee 150 nose fairing (NASA Flight No. 4.12).

4.2.5. Heating During Reentry. Reentry heating is caused by aerodynamic friction and shock effects. The radiation contributions are thus from the heated body and the shocked air.

As a body enters the earth's atmosphere, considerable energy is released through the interaction of the body and the atmosphere. The optical radiations released are functions of the original condition of the event, *e.g.*, velocity, mass, material, angle of attack, and vehicle shape.

Below 70 or 80 km, the atmosphere may be considered to be a continuum, and fluid-dynamics principles may be applied.

A blunt body has a comparatively large normal shock, with the result that considerable air is heated and dissociated. Much of the energy stored during dissociation is not released until the air has passed into the wake. A slender vehicle, on the other hand, has little or no normal shock, so that there is less heating and little dissociation of the air. The frictional drag of the surface does cause the air in the boundary layer to decelerate, but the energy thus made available can in large part be efficiently conducted to the body.

The radiation from the surface may be calculated if the temperature and spectral emissivity of the surface are known. Empirical relationships have been developed which permit one to calculate the heat input for spheres, cones, cylinders, and combinations thereof. These relationships depend upon vehicle velocity, ambient density, and of course the shape of the vehicle. The temperatures reached are functions of the thermal properties of the material of the body. However, the contributions from contaminants in the wake often completely overshadow all other contributions.

4.3. Surface Targets

Infrared surface targets include a wide variety of radiation sources which can be categorized in a number of different ways. There are land targets and sea targets; there are strategic targets and tactical targets; there are fixed targets and mobile targets; there are active targets and passive targets. The wide differences between types of surface targets are evident from a listing of just a few examples: rivers, lakes, woods, fields, roads, railroads, airstrips, buildings, bridges, factories, shipping facilities, power plants, ships, submarine wakes, vehicles, and personnel. In spite of the diversity, it is apparent from this list that most surface targets are opaque, or nearly opaque, bodies and that many of them (though certainly not all) are at close to ambient temperatures (approximately 300°K). While some surface targets (*e.g.*, the exhaust manifolds of tanks or other vehicles, power plants, or blast furnaces) can be distinguished by temperatures which are definitely higher than any of the surroundings, many passive objects near ambient temperature must be recognized by other characteristics, such as shape, size, position, and contrast, in the spatial-distribution pattern of radiance.

4.4. Radiometric Analysis and Discussion

See Chapter 2 (Radiometry), Chapter 6 (Atmospheric Phenomena), and particularly Chapter 5 (Backgrounds) for more complete discussions of these topics. Note that atmospheric phenomena enter into the situation not only because of atmospheric attenuation of source radiation but also through the meteorological effects on the surface temperature and emissivity of radiation sources. This is strikingly illustrated by the "washout effect" [12,13]. Chapter 5 is particularly pertinent because, in general, there is no inherent difference between a background source and a target source of radiation. The designation as one or the other reflects only the interest of the moment – the source of interest is a target; other sources, from which radiation is

received along with target radiation, are the background. Also, for most applications the important considerations are those of target-background contrast, not just of target radiation alone.

The radiance N of an opaque body (which includes most surface targets) consists of that due to self-emission N_e and that due to reflection or scattering of incident radiation N_r:

$$N = N_e + N_r \qquad \text{w cm}^{-2}\ \text{sr}^{-1} \tag{4-1}$$

The relative magnitudes of N_e and N_r depend upon a number of factors, of which the two most important are: (1) the surface properties of the body (including the surface temperature) and (2) the incident radiation. N_e will be more likely to predominate over N_r when the surface has high emissivity and low reflectance (rough and dark), and conversely N_r will tend to be greater when the surface has high reflectance and low emissivity (bright and shiny). The interrelationship is examined in more detail in Section 4.4.3. Considerations of specular *vs* diffuse reflectance and the relative spatial positions of illuminating source, target, and detector also may greatly affect the relative magnitude of N_r which, of course, depends directly upon the incident radiation, especially its magnitude and its distribution in wavelength and in direction.

At night, self-emission N_e usually predominates for military targets in the field and N_r may be neglected, except when the target is irradiated by the source of an active or semiactive detection system (usually operating at wavelengths of 1 μ or less). In sunlight the relative importance of N_e and N_r depends greatly on the wavelength range as well as on target surface conditions (including temperatures). At longer wavelengths, greater than about 4 or 5 μ, reflected or scattered sunlight is relatively unimportant and N_r may usually be neglected; at shorter wavelengths, less than about 1 or 2 μ, self-emission becomes unimportant except for very hot targets, such as an exposed red-hot exhaust pipe, so that for bodies close to ambient temperatures N_e may be ignored. At intermediate wavelengths, from about 1 to 5 μ, either or both N_e and N_r may be important. Reference [14] is an annotated bibliography on emittance and reflectance in the infrared, listing 910 pertinent references.

4.4.1. Thermal Emission. The simplest, and probably the most frequently used, approach to the evaluation of N_e is to consider an opaque solid as a graybody which, at a uniform surface temperature of T°K and in a direction in which its emissivity is ϵ, will radiate according to the Stefan-Boltzmann law:

$$N_e = \epsilon\sigma T^4/\pi \qquad \text{w cm}^{-2}\ \text{sr}^{-1} \tag{4-2}$$

where $\sigma = 5.67 \times 10^{-12}$ w cm^{-2} (°K)$^{-4}$.

If the surface is rough or weathered, so that it approximates a perfectly diffusing or "Lambert law" surface, the emissivity ϵ, and hence N_e, will be the same for all directions, but in general it is a function at least of the angle θ from the normal to the surface, $\epsilon = \epsilon(\theta)$. If the surface is not uniform, it may also vary with azimuth direction φ and with position. The exact relationships are summarized in Section 4.4.3. Nevertheless, the Lambert law assumption is often made for lack of any data to establish the variations of ϵ with respect to direction and/or position.

If ϵ is constant with respect to both direction and wavelength, and the emitting surface is at a uniform temperature T, the radiant intensity of the entire target in a given direction is

$$J_e = N_e \int \cos\theta\, dA = N_e A_p \qquad \text{w sr}^{-1} \tag{4-3}$$

where $A_p \equiv \int \cos\theta \, dA$ is the projected area of the target perpendicular to the given direction (θ is the angle between that direction and the normal to the surface element dA and the integration is carried out over all of the "exposed" surface of the target, *i.e.*, all that can be "seen" from the given direction). If, however, there are temperature gradients, and different areas of the exposed surface are at different temperatures, the radiant intensity is given instead by

$$J_e = \int N_e \cos\theta \, dA \qquad \text{w sr}^{-1} \tag{4-4}$$

where N_e is expressed as a function of the temperature T of the element dA by Eq. (4.2) and the temperature T, in turn, is expressed as a function of the surface coordinates which define dA. The integration may be replaced by, or approximated by, a summation if the surface can be divided into uniform regions each of which radiates in accordance with Eq. (4-2). In that case the radiant intensity of each exposed region is computed separately by Eq. (4-3) and the results are simply added to obtain the radiant intensity of the entire target.

It is only when the target surface has constant emissivity and is also everywhere at uniform temperature that the value of its radiance N_e, or its radiant intensity J_e, as measured with a nonselective radiometer (with equal response at all wavelengths), can be unambiguously associated with its temperature. It is for this reason that the WGIRB (Working Group on Infrared Backgrounds) recommended strongly against the unfortunately all-too-common practice of using temperature as a radiometric unit in lieu of radiance [15].

Even when temperature gradients and differences exist and/or the emissivity is not constant with location and direction (but *is* constant with respect to wavelength), the radiance of each portion (possibly infinitesimal) of the surface of an opaque body is directly associated with the surface temperature of that portion by Eq. (4-2). It is then possible to evaluate or estimate the contrast between different areas with slightly different surface temperatures and/or emissivities by the relation

$$\frac{dN_e}{N_e} = \frac{\sigma(4\epsilon T^3 \, dT + T^4 \, d\epsilon)/\pi}{\epsilon\sigma T^4/\pi}$$

$$= 4dT/T + d\epsilon/\epsilon \qquad \text{dimensionless} \tag{4-5}$$

Larger differences are evaluated by directly computing N_e for each portion by means of Eq. (4-2) and then taking the difference or making the comparison. Contrast is then measured, by analogy with the expression dN/N, by

$$C = (N_1 - N_2)/N = 2(N_1 - N_2)/(N_1 + N_2) \tag{4-6}$$

If the background radiance N_B is regarded as the reference level and one wishes to designate the contrast, with respect to this background level, of a target of radiance N_T, one may instead compute the contrast as [16]

$$C = (N_T - N_B)/N_B \qquad \text{dimensionless} \tag{4-7}$$

It must be strongly reemphasized that the foregoing relations, in terms of *total* radiation (all wavelengths) have all been based on the graybody (constant-spectral-emissivity) assumption, which is usually only approximately true of real targets. Furthermore, application of these relations also implies a detector response to total radiation, *i.e.*, nonselective response or constant spectral responsivity. Spectral distribution has been ignored, although the strong wavelength dependence of atmospheric attenuation and the frequent use of spectral filters and spectrally selective

detectors makes it a very important consideration in many, if not most, applications. Accordingly, these relations will ordinarily give only approximate or qualitative results for real situations, and often the approximation may be very poor. On the other hand, any evaluation of spectral effects as functions of source temperature, based on the Planck law (see Chapter 2), is best carried out in terms of each particular situation. Attempts to write expressions which will apply generally in a wide variety of circumstances rapidly become prohibitively complex.

4.4.2. Reflection. The measurement and specification of reflectance, even for opaque materials where multiple internal reflections are not involved, is not at all a simple matter. The additional complications which arise with semitransparent materials are well discussed in [17] and [18], and will not be considered further here.

It is particularly important here to recognize that the value of total reflectance, ρ, which is related by Kirchhoff's law to the absorptance, α, or the emissivity, ϵ, of an opaque body,

$$\alpha = \epsilon = (1 - \rho) \qquad \text{dimensionless} \tag{4-8}$$

may be quite different from the value of directional or partial reflectance (see Section 2.10), which, together with the spatial distribution of incident radiation and the locations of the target and detector, determines the value of N_r in Eq. (4-1). The basic relationships involved are clearly and concisely summarized in an appendix to a paper by Richmond beginning on page 151 of [19].

In general, measured values of reflectance may be applied correctly only to situations which involve the same geometric relations between the source of irradiance, the reflecting surface, and the receiver or detector. Any change in the geometry of either the incident beam or the reflected beam of radiation may result in, or require, a different value of reflectance.

Possible confusion in terminology should be noted. Throughout most of [19], a distinction is made between *reflectivity,* which is defined as the property of a material (*i.e.,* measured with an ideally smooth and clean surface), and *reflectance,* which is the fraction of incident radiation reflected from a particular sample, regardless of its surface condition. The terms *emissivity* and *emittance* are used similarly with corresponding meanings, and the power per unit area emitted by a source, which we have called the *radiant emittance, W* (following [20]) is instead called the *emissive power.* Still another term, *albedo,* is frequently used by astronomers and meteorologists to designate the total radiant reflectance of natural objects. Similarly, *visual albedo* is used to designate the luminous reflectance [21] (see Chapter 2 for further comments on this nomenclature).

4.4.3. Factors Affecting the Temperatures of Passive Targets. Since the self-emission N_e of an opaque target is so closely dependent upon its *surface* temperature, it is important to recognize the factors which determine that temperature for a passive or inert object under field conditions. This is a very complex situation that, in practice, is usually not amenable to quantitative treatment except for making very rough approximations [22]. However, qualitatively it is always useful and important to recognize and take into account the various contributing factors, such as the incident radiation or other source of heat, the recent history of incident radiation or other heating, the absorptivity (see Eq. (4-8)), the heat conductivity, the heat capacity, the size and shape and material of the target, and other ambient conditions such as convective cooling or heating by winds or cooling by rain, or other precipitation or condensation (dew), and by its subsequent evaporation [23]. While the foregoing list may not be

exhaustive, it probably includes the most important parameters and is complete enough to suggest others which may be pertinent in a particular case.

The forms in which heat energy reaches the surface of the earth, both from above and from below, are shown in diagrams of the daytime heat balance (Fig. 4-11), and the nighttime heat balance (Fig. 4-12) [24]. These are simplified diagrams of average or gross effects and do not take into account the local gradients which result from differences in heat capacity and/or heat conductivity, interactions between terrain configuration and sun (shadowing) or winds (sheltering), etc. The resulting variations in surface temperatures give rise to most of the observed diurnal and seasonal variations and the effects of overcast histories. A thorough understanding of the phenomena involved is essential for the interpretation of strip maps obtained with infrared scanners and for the development of effective camouflage techniques.

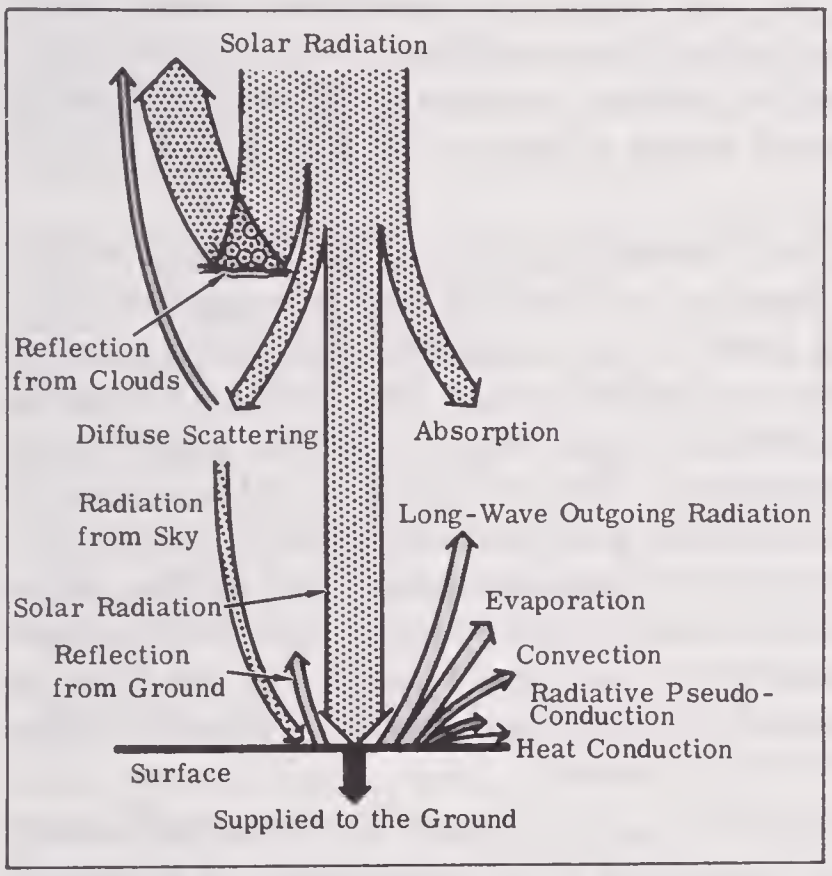


FIG. 4-11. Heat balance at the earth's surface at midday ([24], taken from *The Climate Near the Ground,* by R. Geiger, Harvard University Press, 1950).

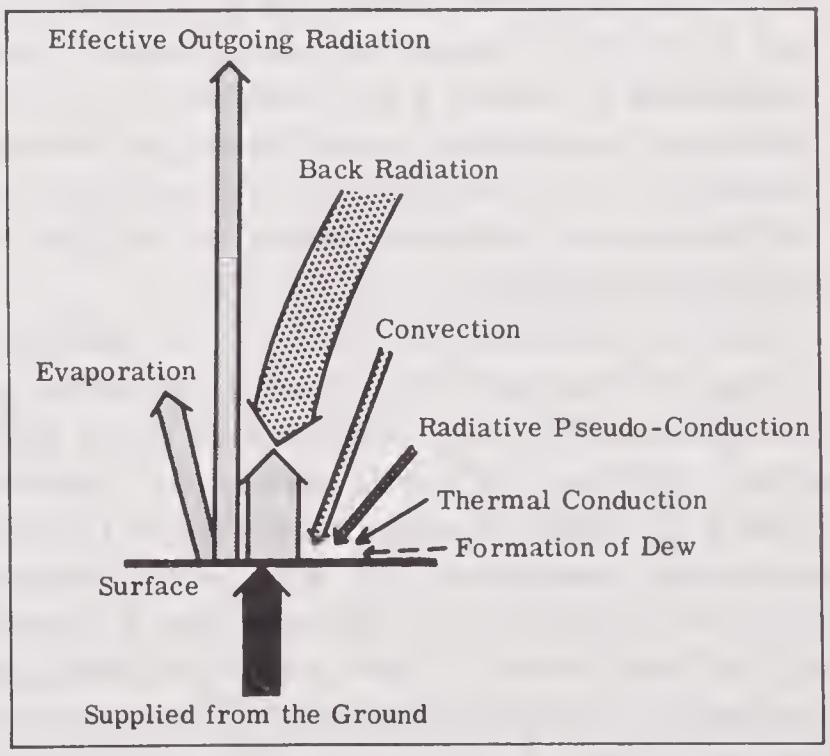


FIG. 4-12. Heat balance at the earth's surface at night ([24], taken from *The Climate Near the Ground,* by R. Geiger, Harvard University Press, 1950).

4.5. Measured Values of Radiant Emissivity and Reflectance

The complexities involved in the description and specification of the interrelated parameters of radiant emissivity and reflectance have been summarized in preceding sections, particularly in Chapter 2. The basic concepts and relations are not always clearly defined or applied in reports of measurement results, and a variety of measuring techniques is employed, with the result that uncertainties exist regarding many published values [25]. Some representative values of emissivities and/or reflectances are presented here, but it must be reemphasized that they can be regarded as accurate and strictly applicable only in situations where the ray geometry corresponds to that used in making the measurements, and also only where the surface conditions (weathering, corrosion, etc.) of the target or sample correspond to those of the sample measured. Usually, however, it is difficult if not impossible even to determine the degree to which these conditions are satisfied or reproduced. Often the pertinent information is not included in the measurement report, but, even when it is, verification is not a simple matter. Consequently, the reported values are often useful only to indicate the orders of magnitude involved and to suggest the probable relationship or contrast between a target and other targets or background objects as detected by a particular infrared device.

It is only possible to present here a few representative curves and tabulated values. Additional measurement reports are found in the literature, and an extensive listing of published material prior to 1959 is found in [14]. Reference [26] contains an extensive tabulation of published experimental results through 1957, giving considerable detail about the coverage of the measurements and references to the measurement reports but not the measured values themselves. More detailed data on the reflectance and emissivity of different materials, particularly of pure substances with clean surfaces, measured in the laboratory (reflectivities and emissivities rather than reflectances and emittances or emissances, in the much-debated terminology for distinguishing material properties from the parameters of particular samples [19]) are found in standard references, such as the *Handbook of Chemistry & Physics* (Chemical Rubber Publishing Co.) and the *International Critical Tables.*

The measurements shown here are all for opaque substances, and the reflectances are assumed to be at least approximately equivalent to total reflectance ρ, or directional reflectance ρ_d, as defined in Chapter 2 (note particularly the distinction between the latter and the partial reflectance or reflectance distribution function ρ') so that they may be related to the corresponding emissivities and absorptances by Eq. (4-8). As emphasized in Section 4.4.3., however, Eq. (4-8) holds strictly only for monochromatic radiation, for radiation consisting only of wavelengths for which values of reflectance, emissivity, and absorptance do not vary with wavelength, or where the reflectance and absorptance values are those for incident radiation with a blackbody (graybody) spectral distribution.

Spectral reflectance curves for a few varieties of ordnance materials are presented in Figs. 4-13 to 4-16 [27]. Similar curves for surfaces and finishes of naval interest are shown in Figs. 4-17 through 4-24 [28]. Reference [29] covers an extensive study of the reflectances of a wide variety of terrain features and of paints and finishes, including the effects of water immersion on the latter. Some of the spectral reflectance curves are presented in Fig. 4-25 as they were summarized in [30], where these and other similar data from [29] were used to compute the emissivities in the 3-5-μ and 8-13-μ bands listed in Table 4-5, which also includes values, obtained similarly from the spectral reflectance curves of [28], for reflectances in the 0.7-1.0-μ band and emissivities in the 1.8-2.7-μ band. The importance of the substrate to which a paint is

applied and the effect that the substrate may have on the spectral reflectance of the painted surface is illustrated by Figs. 4-26 and 4-27 [29].

Curves of directional reflectance for some paving and roofing materials and paints, showing the effects of different angles of incidence in different spectral regions, appear in Figs. 4-28 through 4-33 [17]. Note that the "directional reflectivity," $r_{\alpha\beta}$, of [17] (plotted in these figures) is related to the partial reflectance ρ', as defined in Eq. (2-45), as follows:

$$r_{\alpha\beta} = \rho'(\alpha, \phi, \beta, \phi \pm \pi) \cos \alpha \cos \beta \tag{4-9}$$

TABLE 4-5. REFLECTANCE (ρ) AND EMISSIVITY (ϵ) OF COMMON TERRAIN FEATURES*

	0.7-1.0 μ	1.8-2.7 μ	3-5 μ	8-13 μ
Green Mountain Laurel	$\rho = 0.44$	$\epsilon = 0.84$	$\epsilon = 0.90$	$\epsilon = 0.92$
Young Willow Leaf (dry, top)	0.46	0.82	0.94	0.96
Holly Leaf (dry, top)	0.44	0.72	0.90	0.90
Holly Leaf (dry, bottom)	0.42	0.64	0.86	0.94
Pressed Dormant Maple Leaf (dry, top)	0.53	0.58	0.87	0.92
Green Leaf Winter Color – Oak Leaf (dry, top)	0.43	0.67	0.90	0.92
Green Coniferous Twigs (Jack Pine)	0.30	0.86	0.96	0.97
Grass – Meadow Fescue (dry)	0.41	0.62	0.82	0.88
Sand – Hainamanu Silt Loam – Hawaii	0.15	0.82	0.84	0.94
Sand – Barnes Fine Silt Loam – So. Dakota	0.21	0.58	0.78	0.93
Sand – Gooah Fine Silt Loam – Oregon	0.39	0.54	0.80	0.98
Sand – Vereiniging – Africa	0.43	0.56	0.82	0.94
Sand – Maury Silt Loam – Tennessee	0.43	0.56	0.74	0.95
Sand – Dublin Clay Loam – California	0.42	0.54	0.88	0.97
Sand – Pullman Loam – New Mexico	0.37	0.62	0.78	0.93
Sand – Grady Silt Loam – Georgia	0.11	0.58	0.85	0.94
Sand – Colts Neck Loam – New Jersey	0.28	0.67	0.90	0.94
Sand – Mesita Negra – lower test site	0.38	0.70	0.75	0.92
Bark – Northern Red Oak	0.23	0.78	0.90	0.96
Bark – Northern American Jack Pine	0.18	0.69	0.88	0.97
Bark – Colorado Spruce	0.22	0.75	0.87	0.94

*Estimated average values of reflectance ρ, or emissivity $\epsilon = 1 - \rho$, in the indicated wavelength bands, read from the spectral reflectance curves of [29] (some of which are shown in Fig. 4-25).

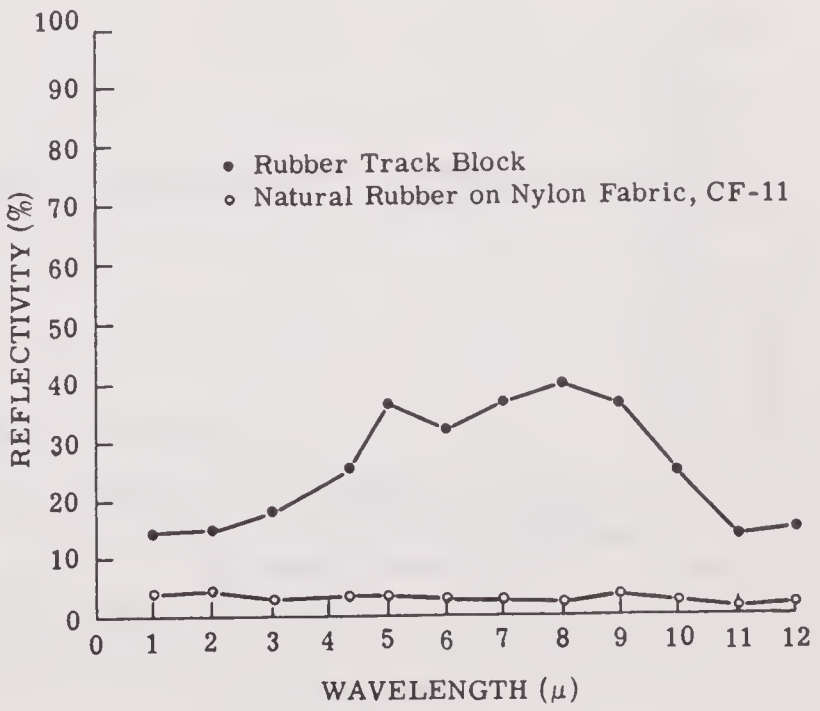


FIG. 4-13. Spectral reflectance of rubber track block, natural rubber on nylon fabric, CF-11 [27].

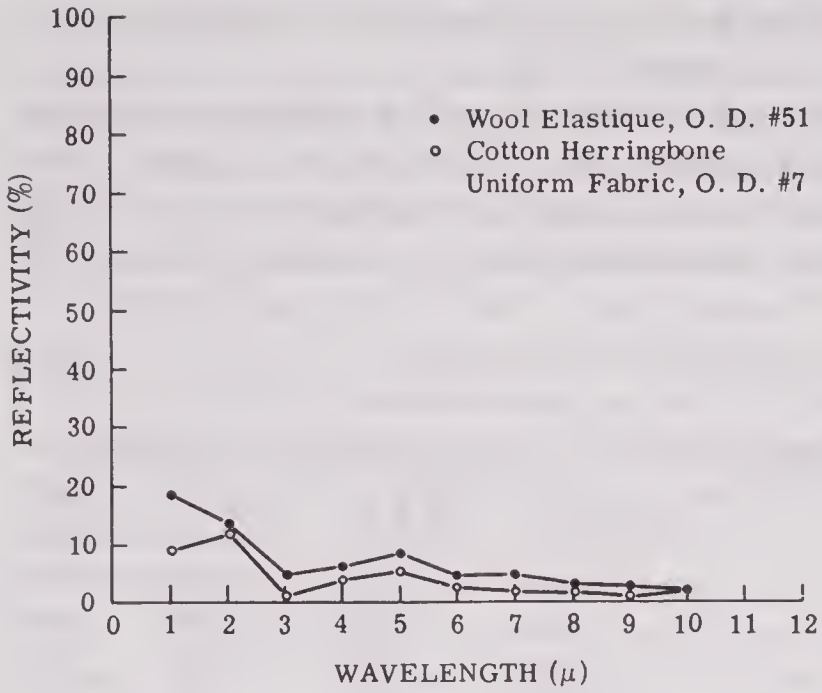


FIG. 4-14. Spectral reflectance of cotton herringbone uniform fabric, D.C. No. 7 Wool Elastique, O.D. No. 51 [27].

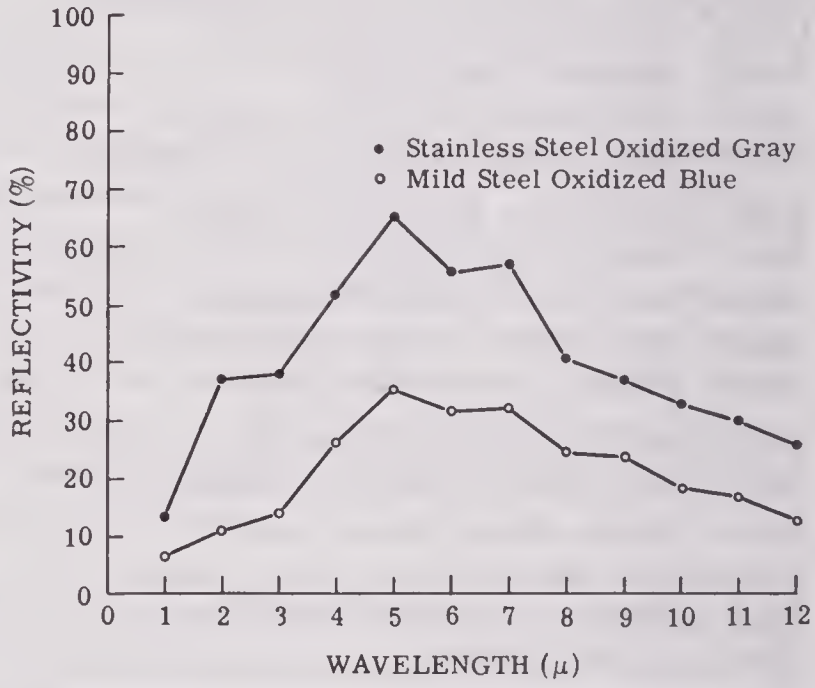


FIG. 4-15. Spectral reflectance of mild steel oxidized blue and stainless steel oxidized grey [27].

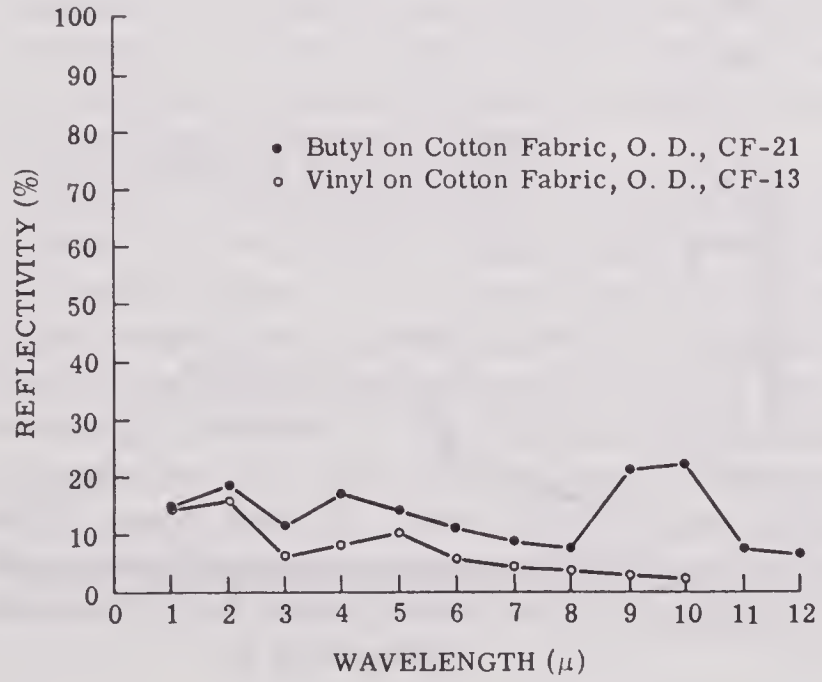
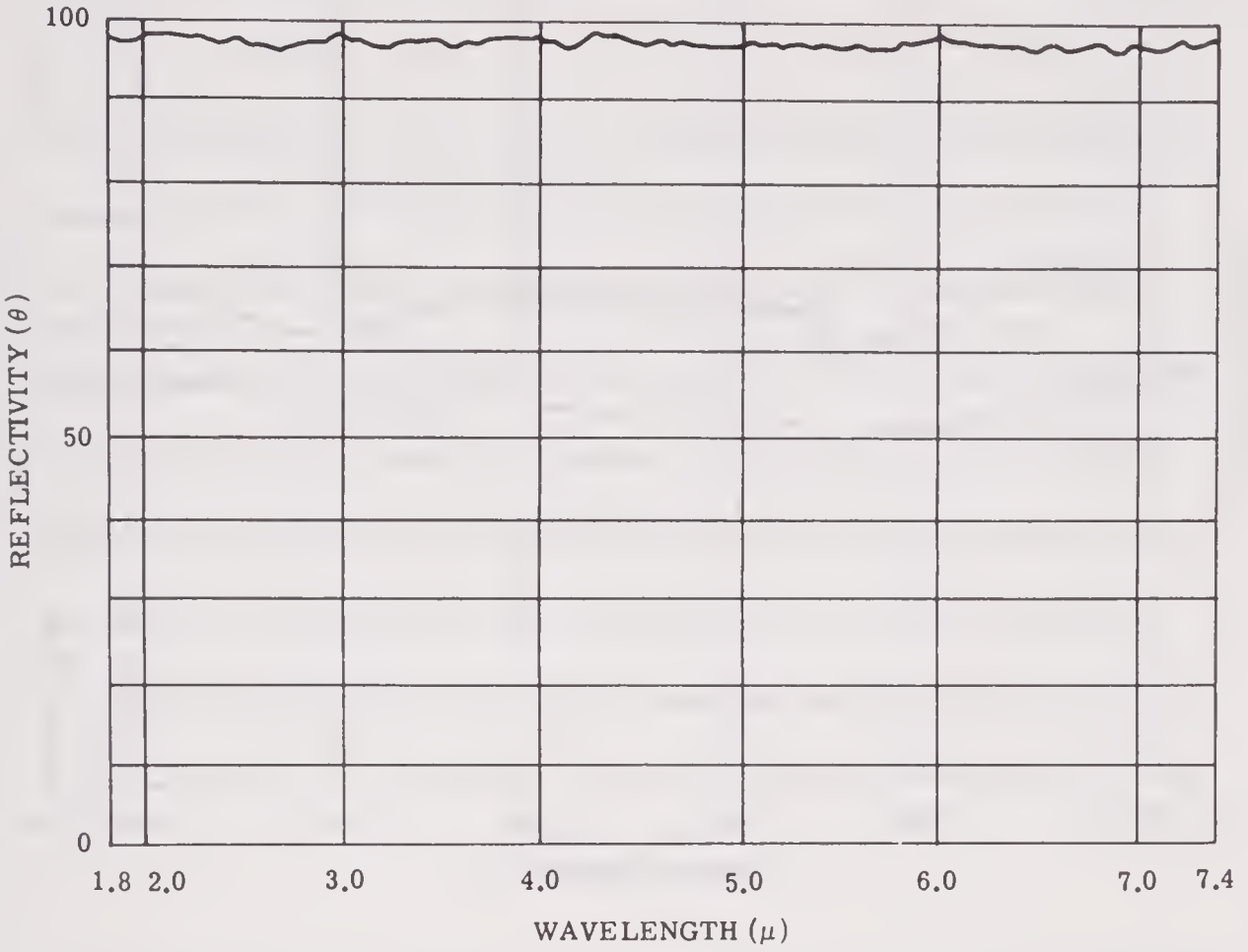


FIG. 4-16. Spectral reflectance of butyl on cotton fabric, O.D., CF-21 and vinyl on cotton fabric, O.D., CF-13 [27].



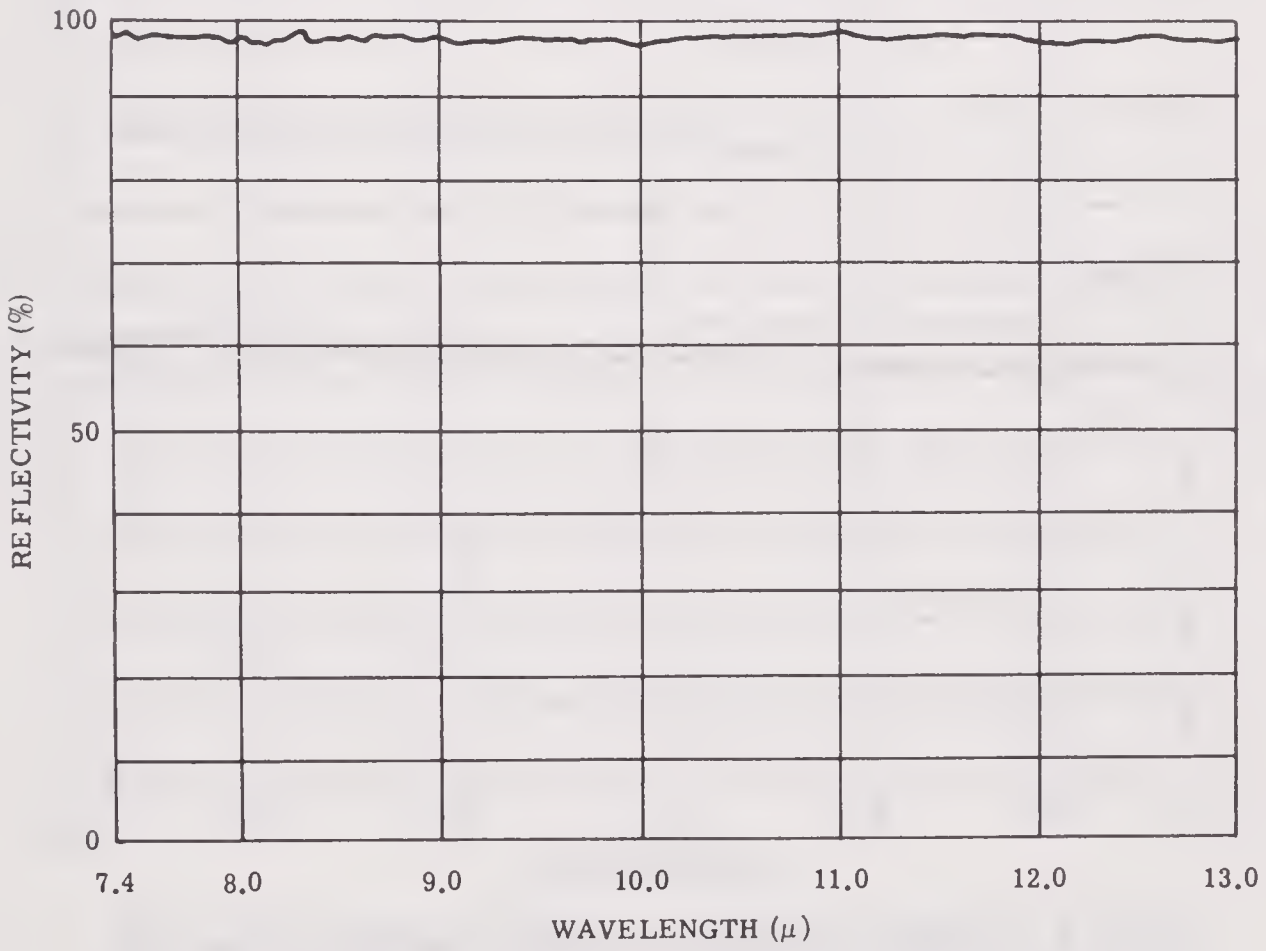
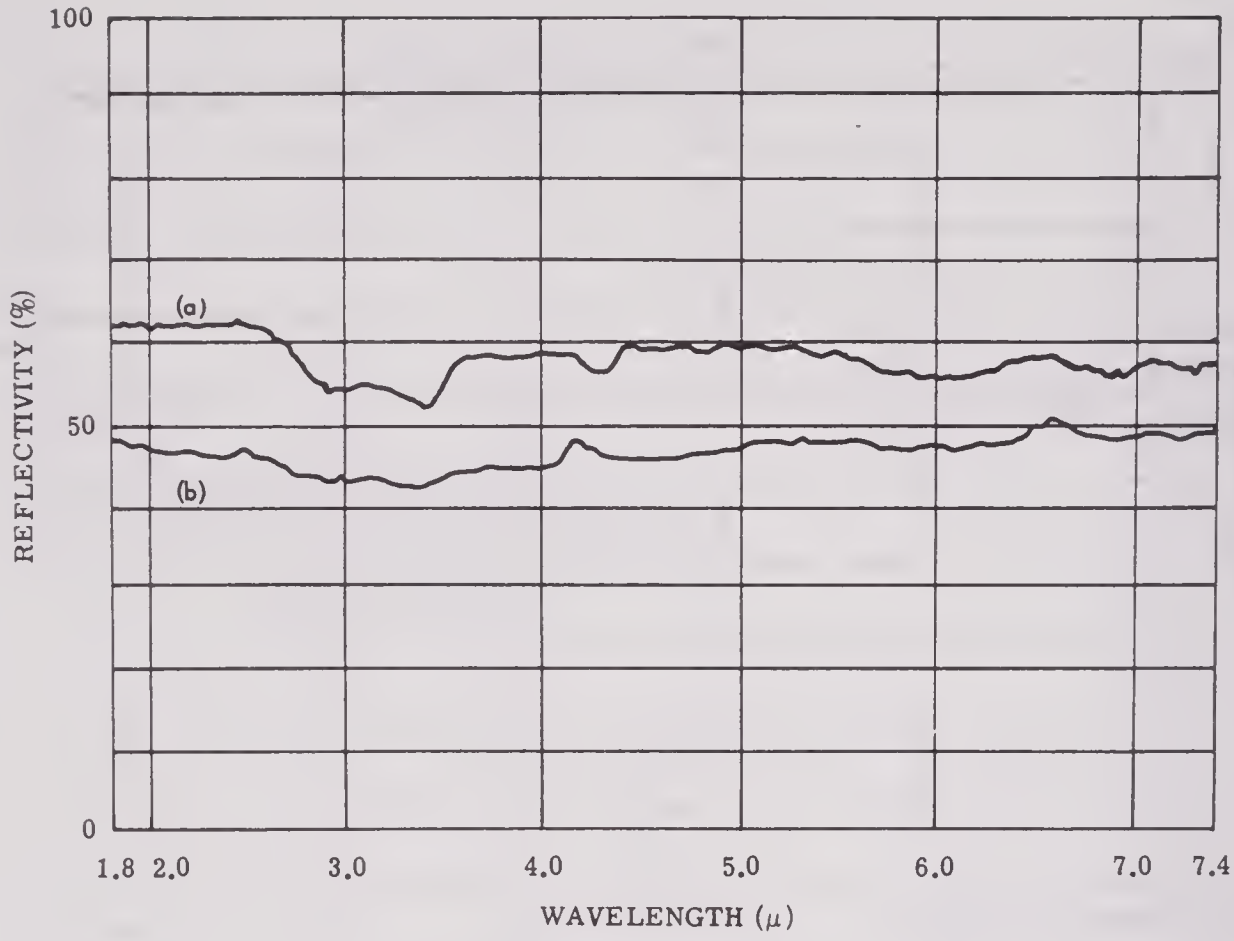


FIG. 4-17. Spectral reflectance of aluminum foil, 0.001 in. RM-216 (Reynolds Metal Co.) [28].



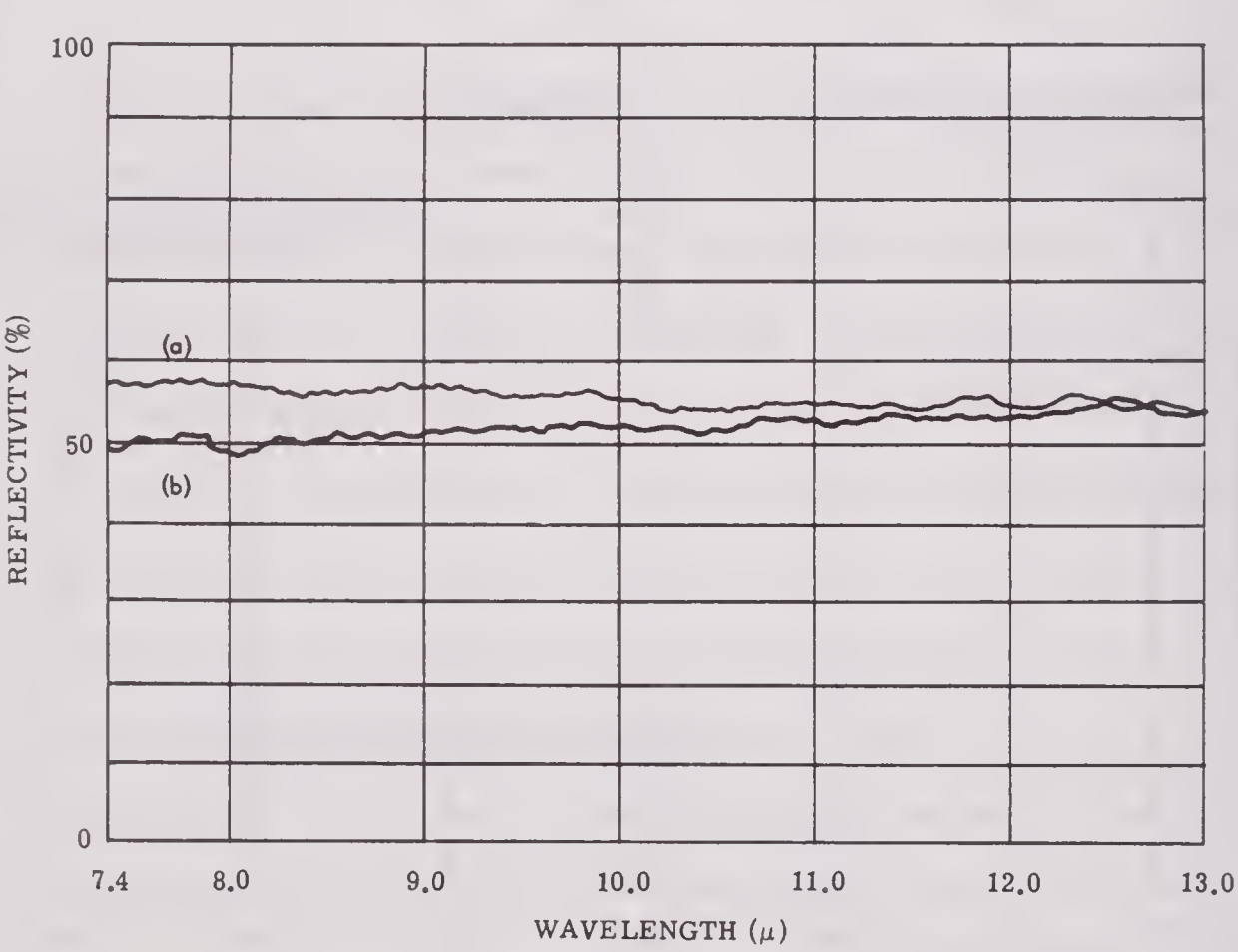


FIG. 4-18. (*a*) Spectral reflectance of aluminum, asphalt base, No. 3483 (Sears Roebuck and Co.), and (*b*) spectral reflectance of aluminum lacquer, No. S-2432-C (Stoner-Mudge Inc.) [28].

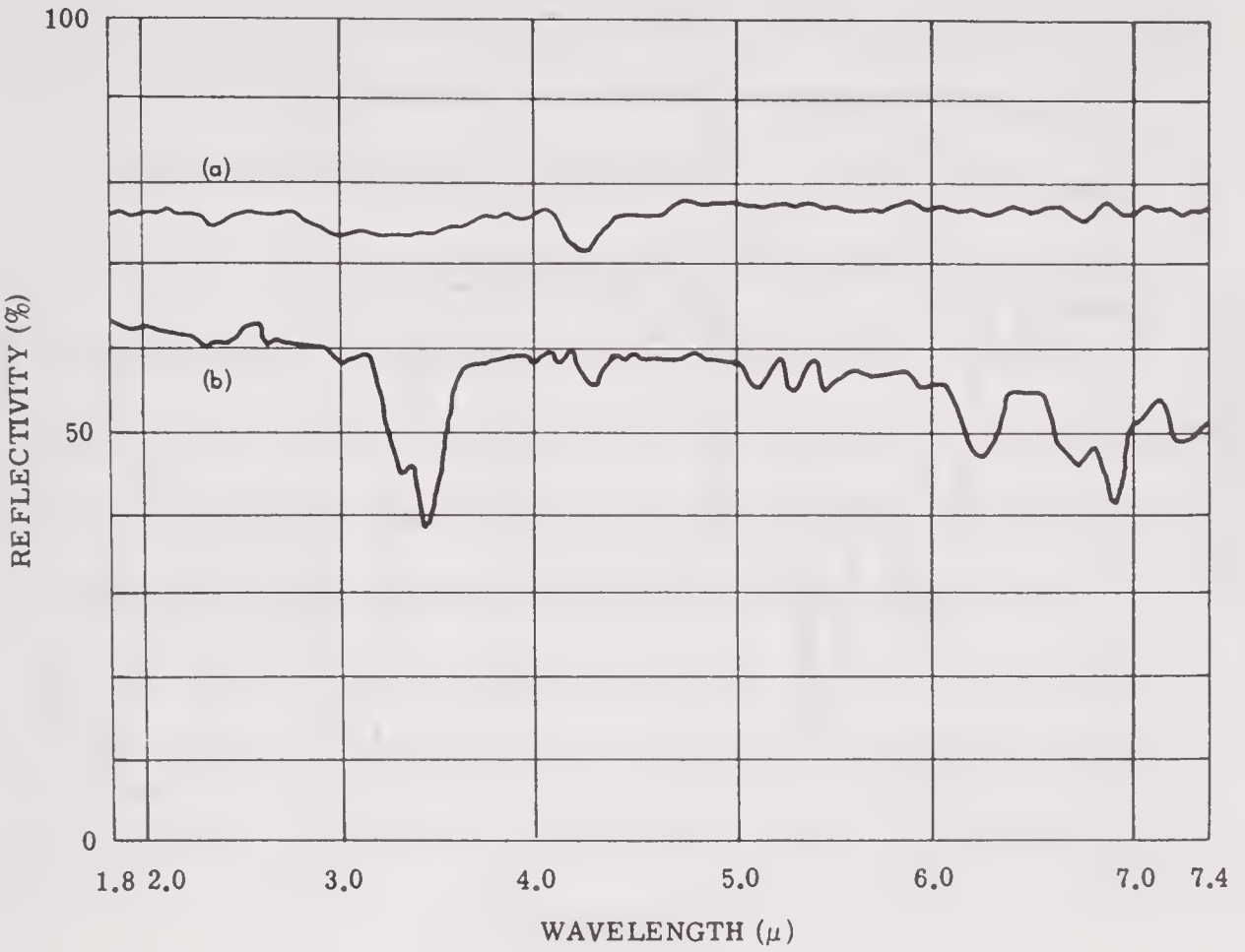
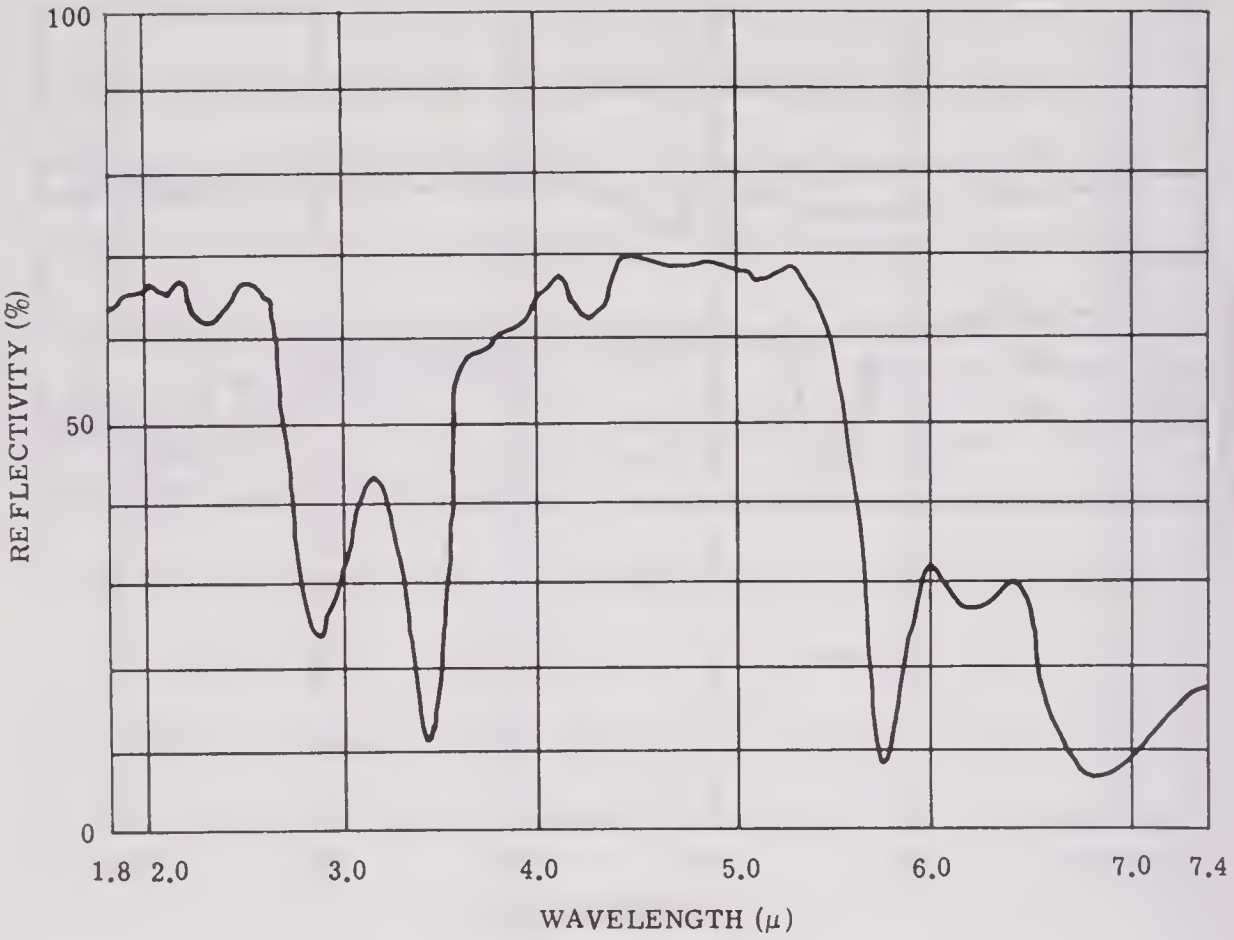


FIG. 4-19. (*a*) Spectral reflectance of Kerpo No. 25 aluminum, and (*b*) Kerpo spray coat, gold (Protective Coatings Corp.) [28].



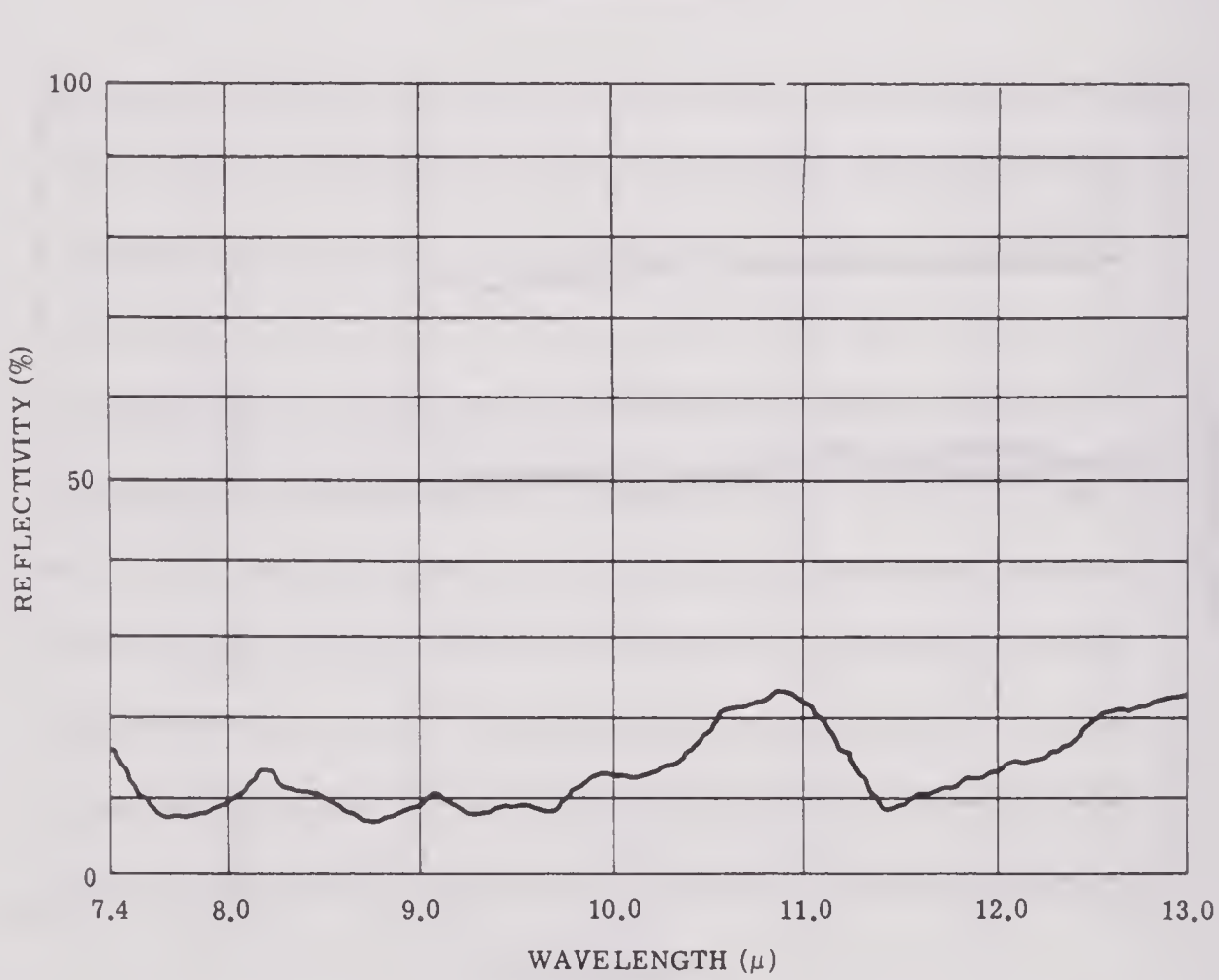


FIG. 4-20. Spectral reflectance of enamel, Chinese red, interior, No. 2816 decoret enamel (W. P. Fuller and Co.) [28].

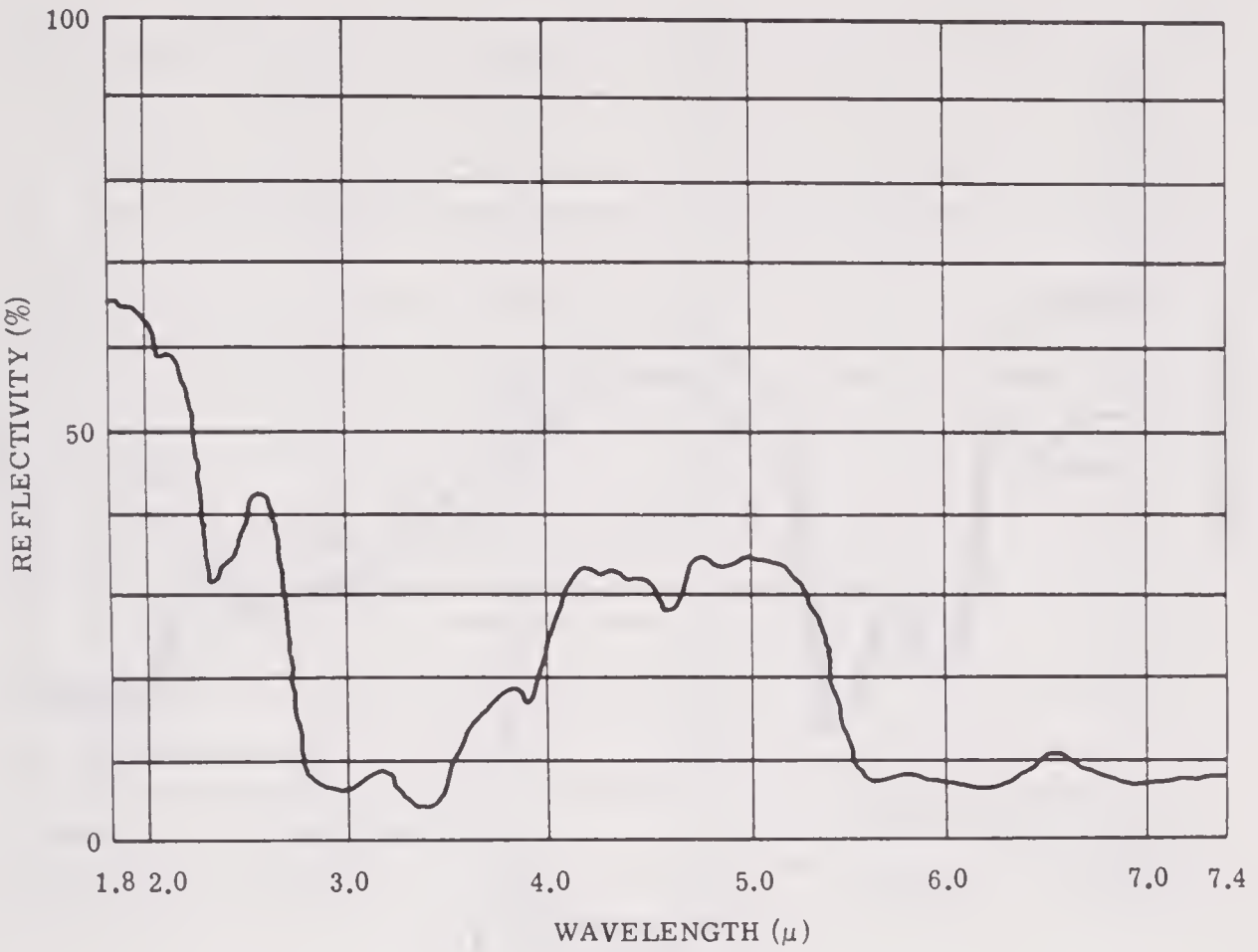


FIG. 4-21. Spectral reflectance of enamel, white, exterior No. 175 (Walter N. Boysen Co.) [28].

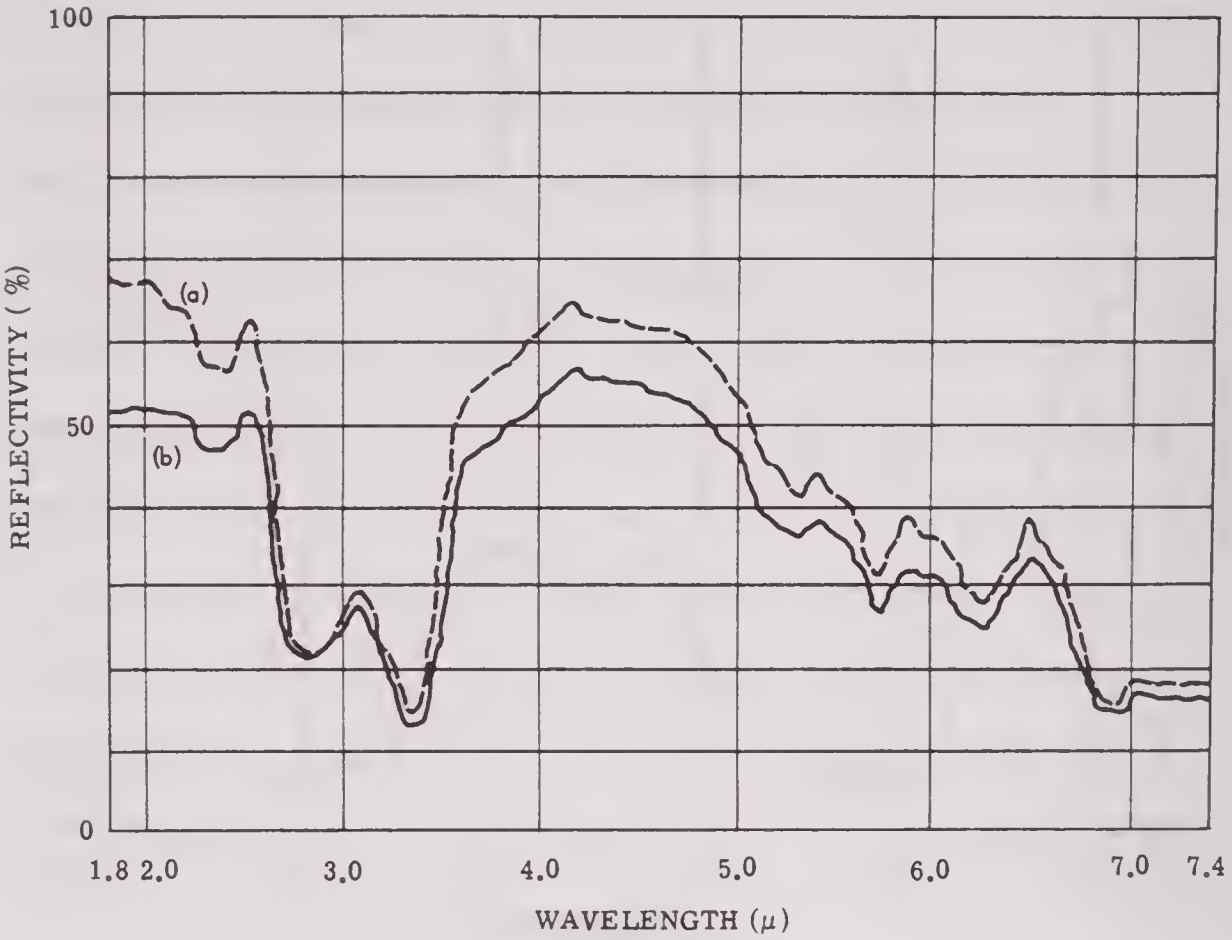
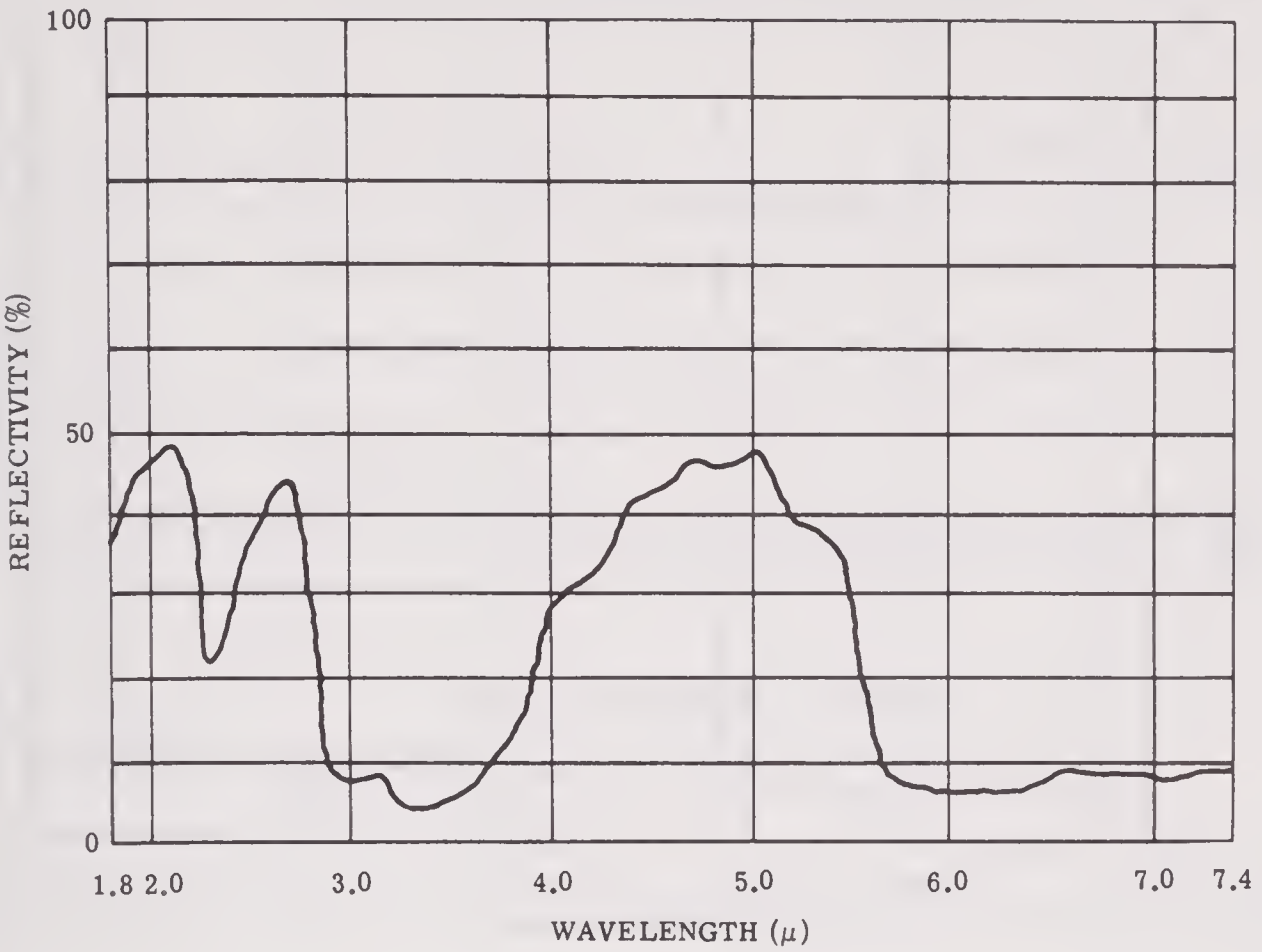


FIG. 4-22. Spectral reflectance of (*a*) ideal masonry, No. 150 (red), (*b*) ideal masonry, No. 170 (green), exterior masonry paint (Ideal Chemical Products Inc.) [28].



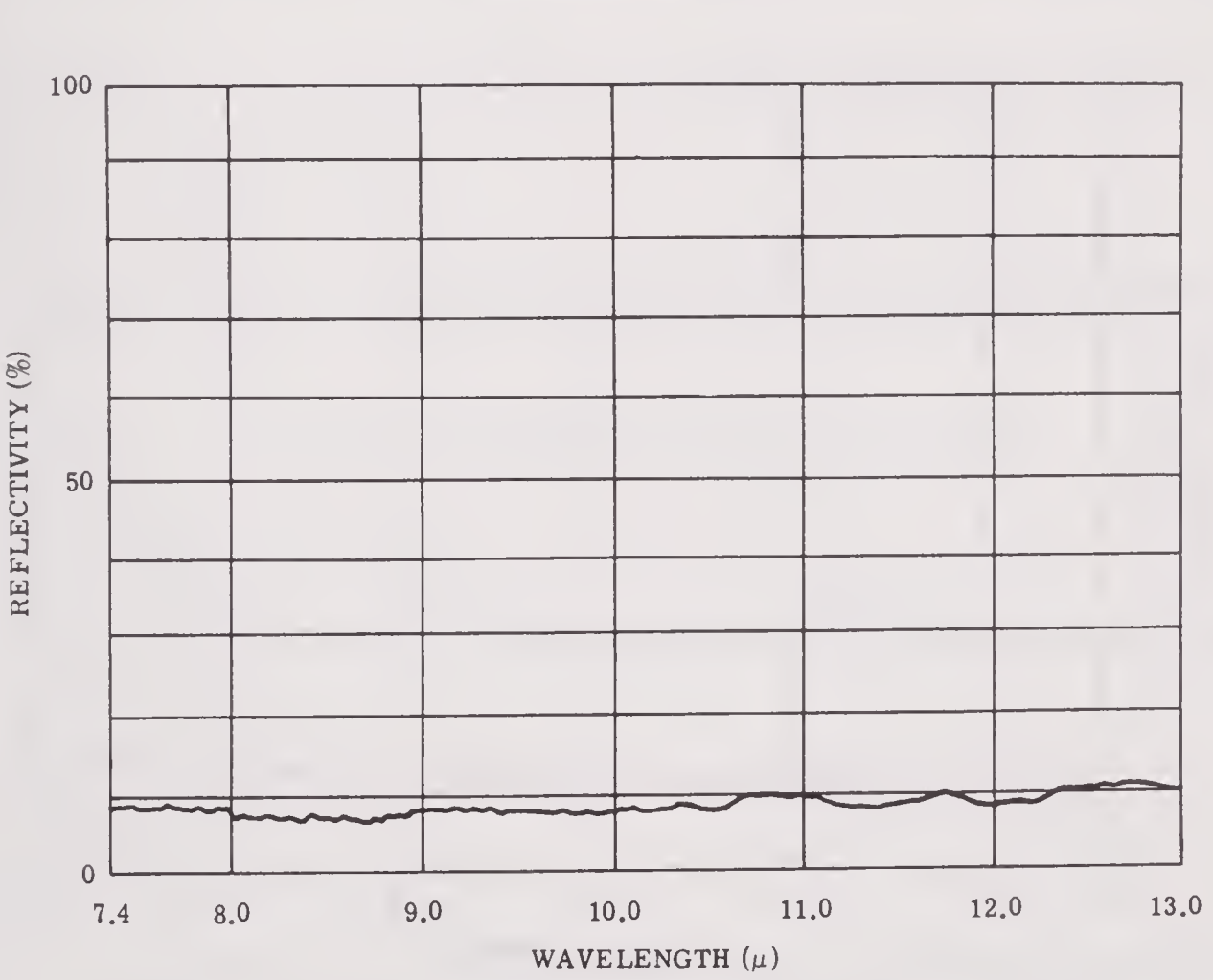
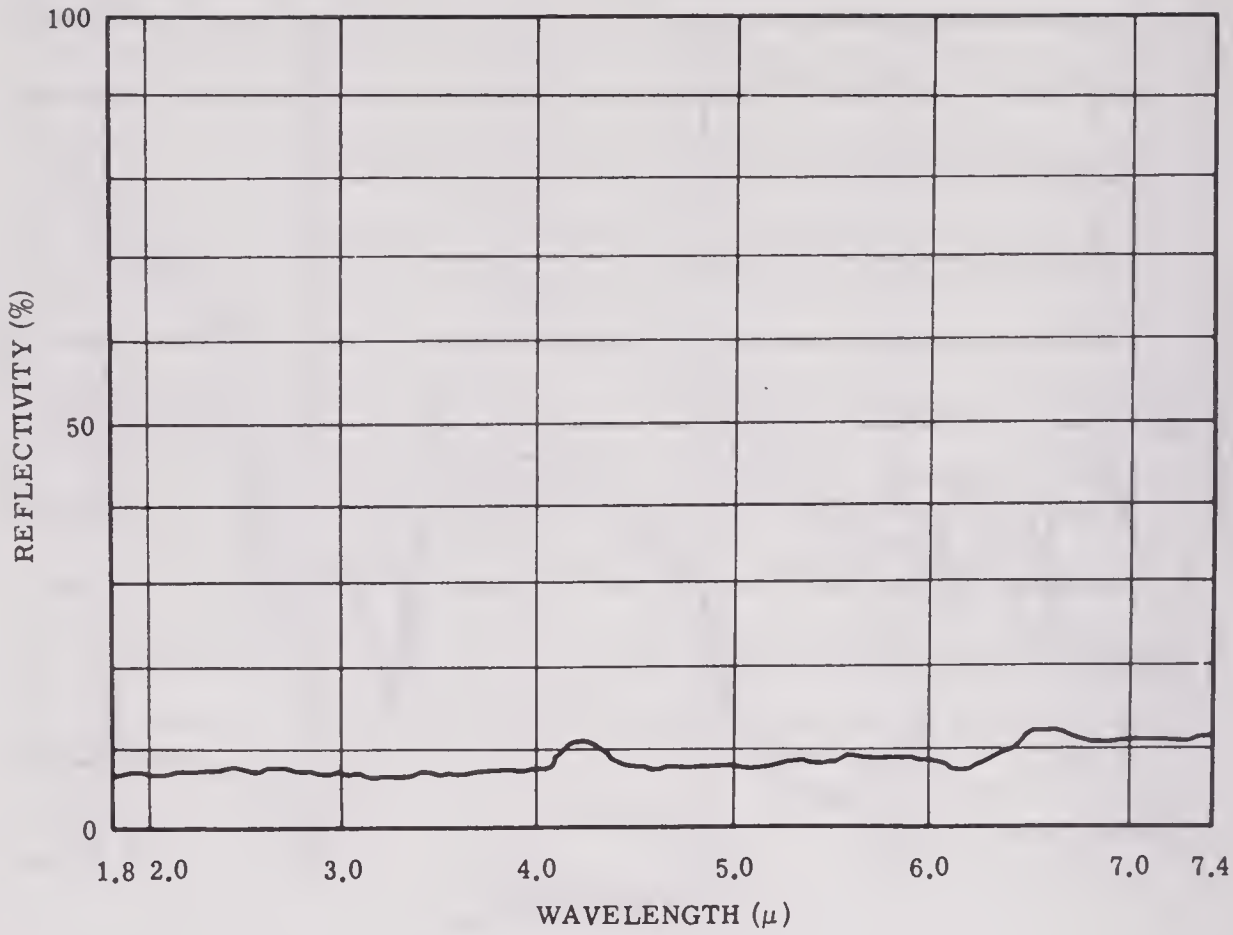


FIG. 4-23. Spectral reflectance of asphaltic road material, SC-4 (Standard Oil Co. of California) [28].



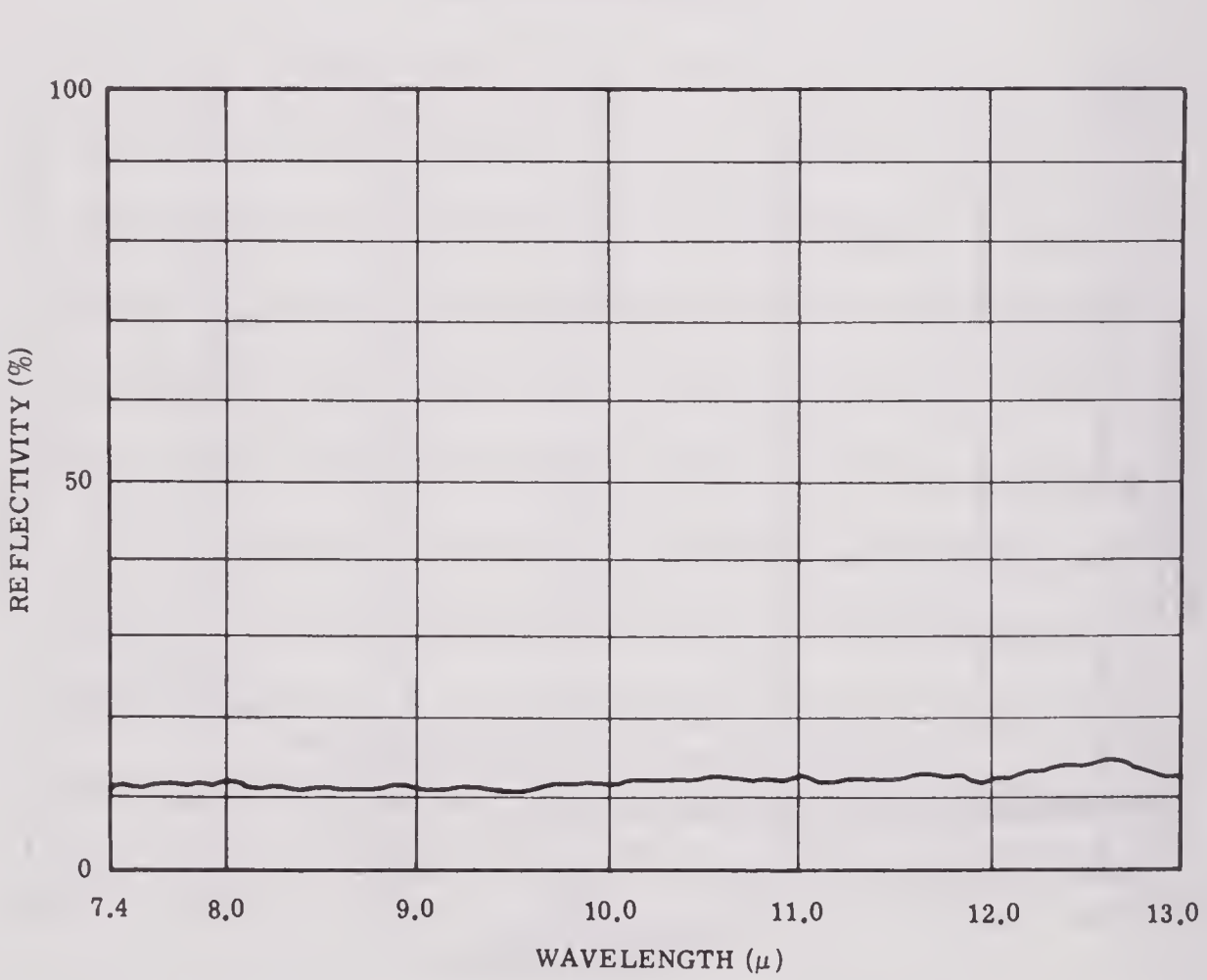


FIG. 4-24. Spectral reflectance of coal tar pitch, melting point 170-180°F (Barrett Div., Allied Chemical and Dye Corp.) [28].

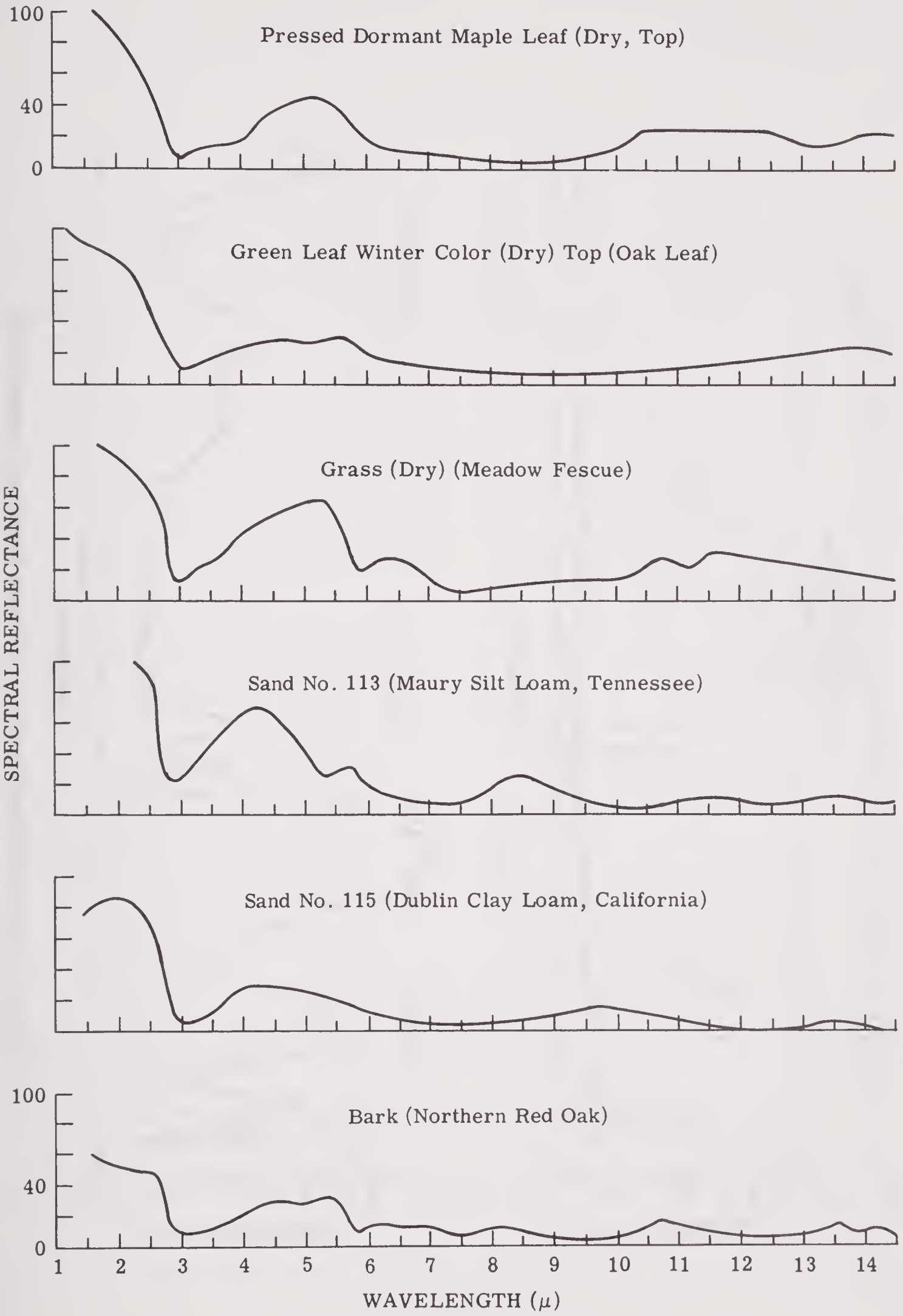


FIG. 4-25. Spectral reflectance of six terrain features in the 1-15-μ region (based on [30]; original source) [29].

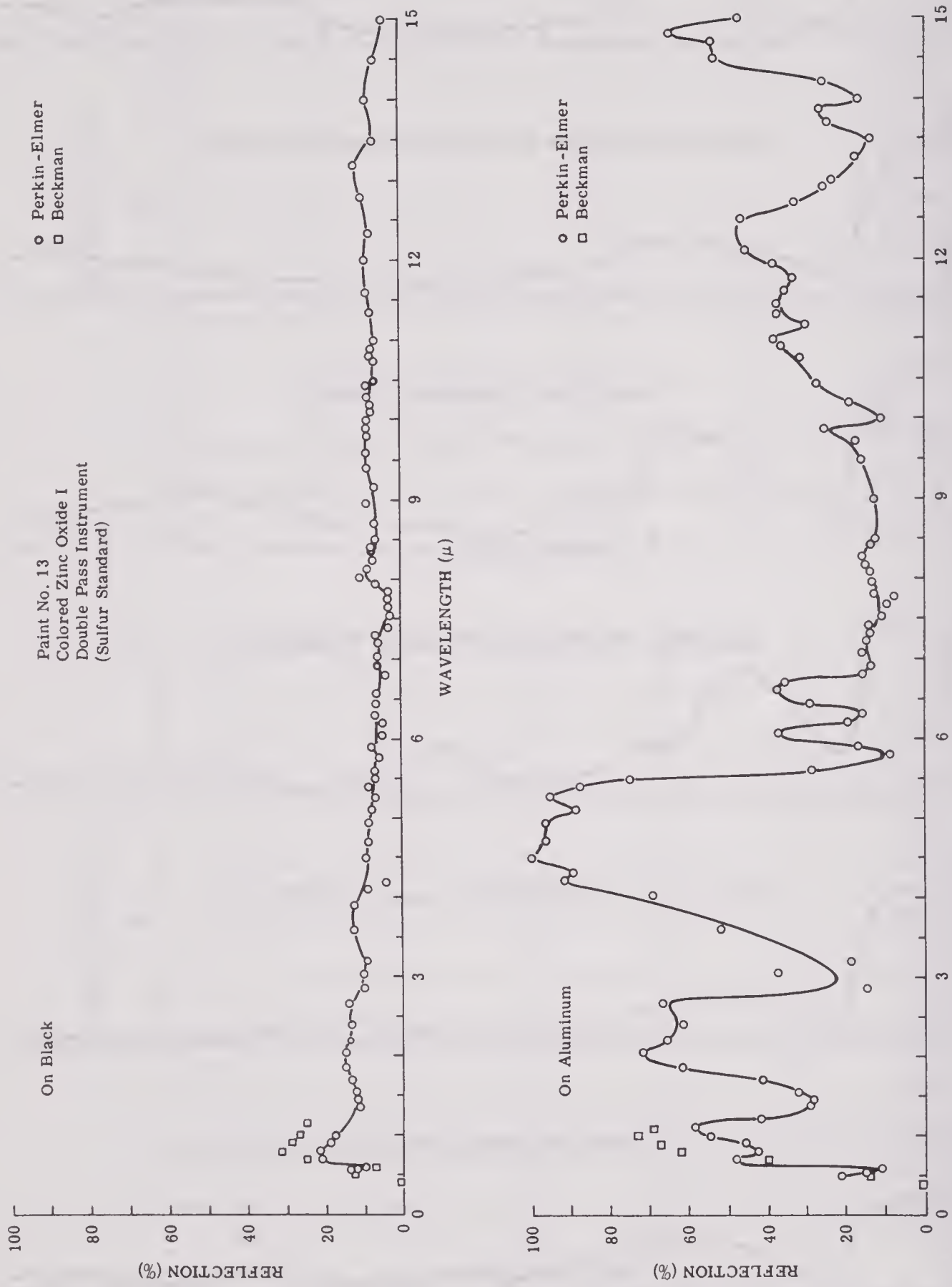


FIG. 4-26. Spectral reflectance of paint No. 13, showing dependence on substrate [29].

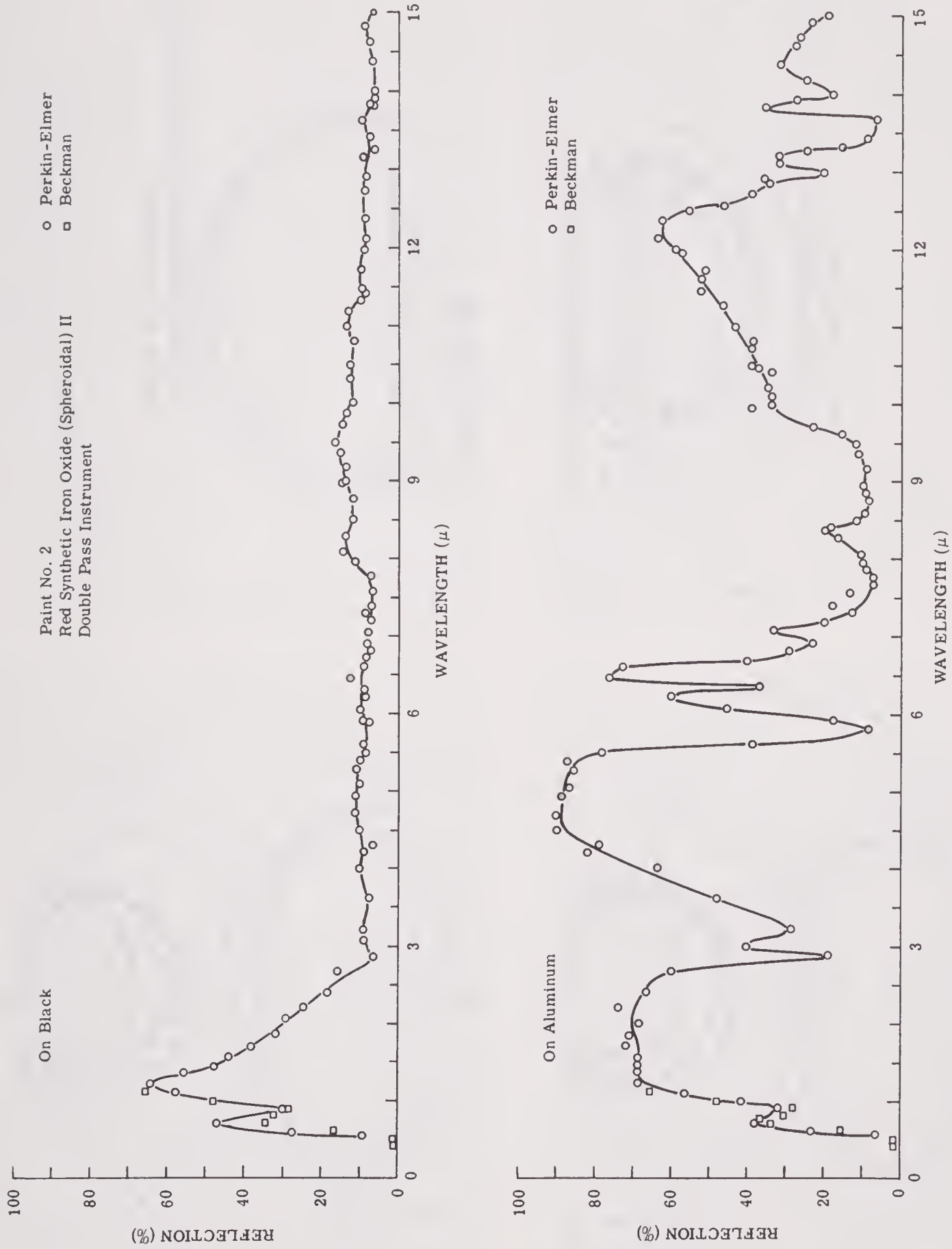
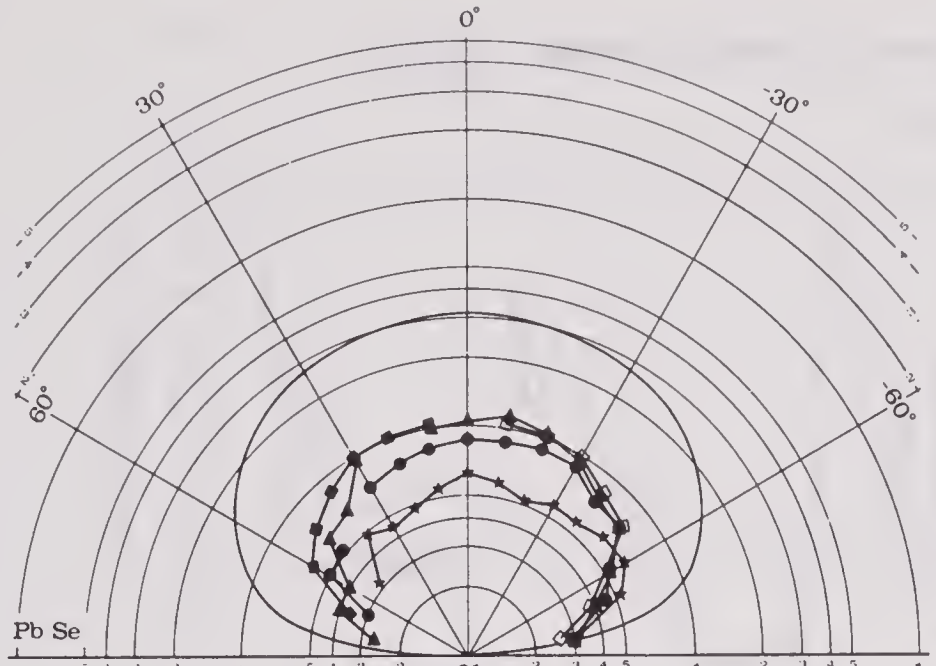
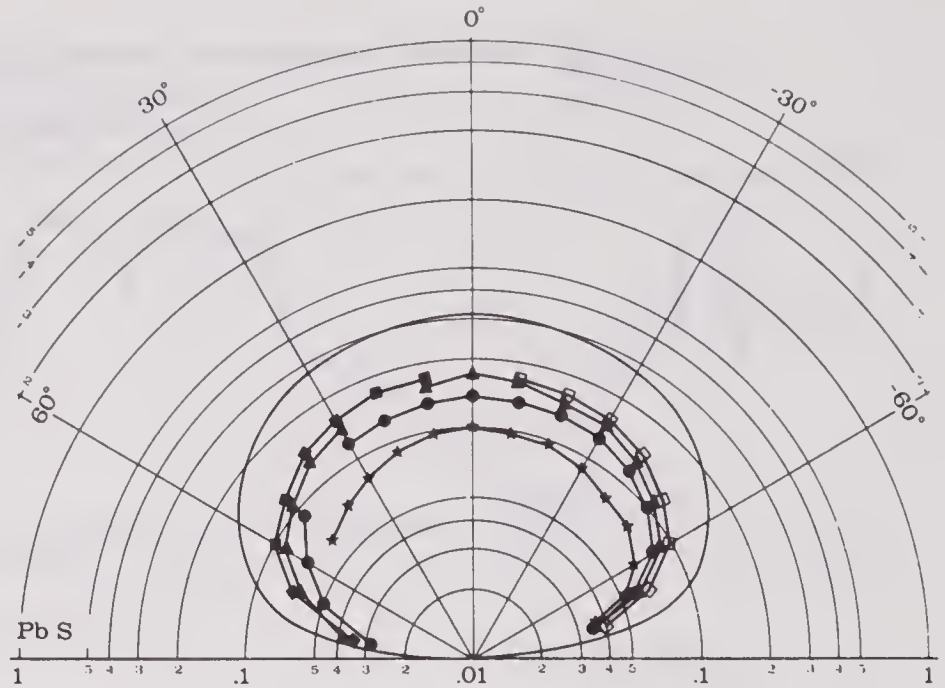
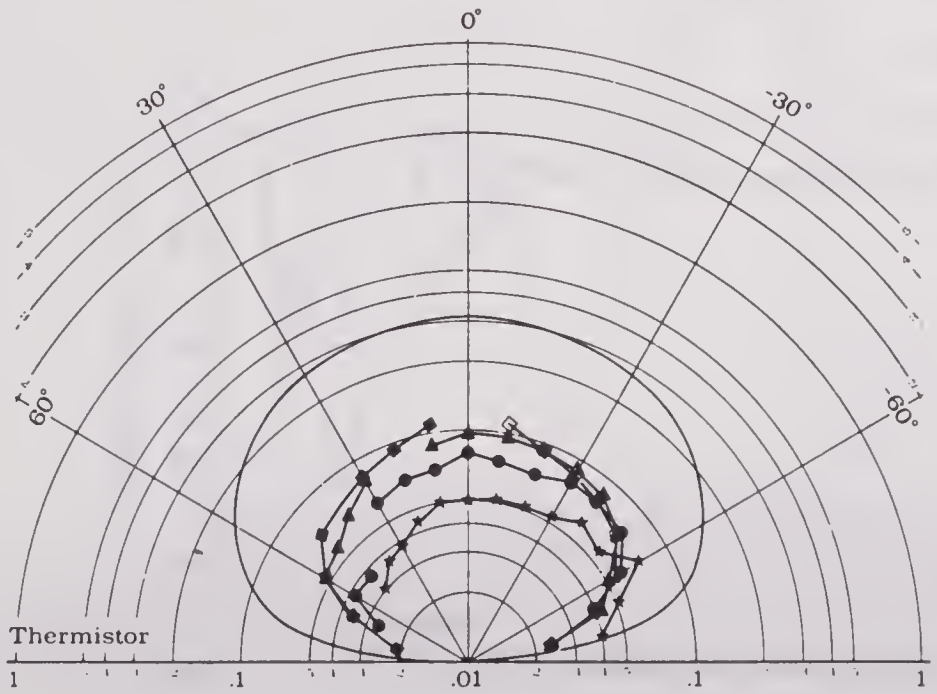


FIG. 4-27. Spectral reflectance of paint No. 2, showing dependence on substrate [29].





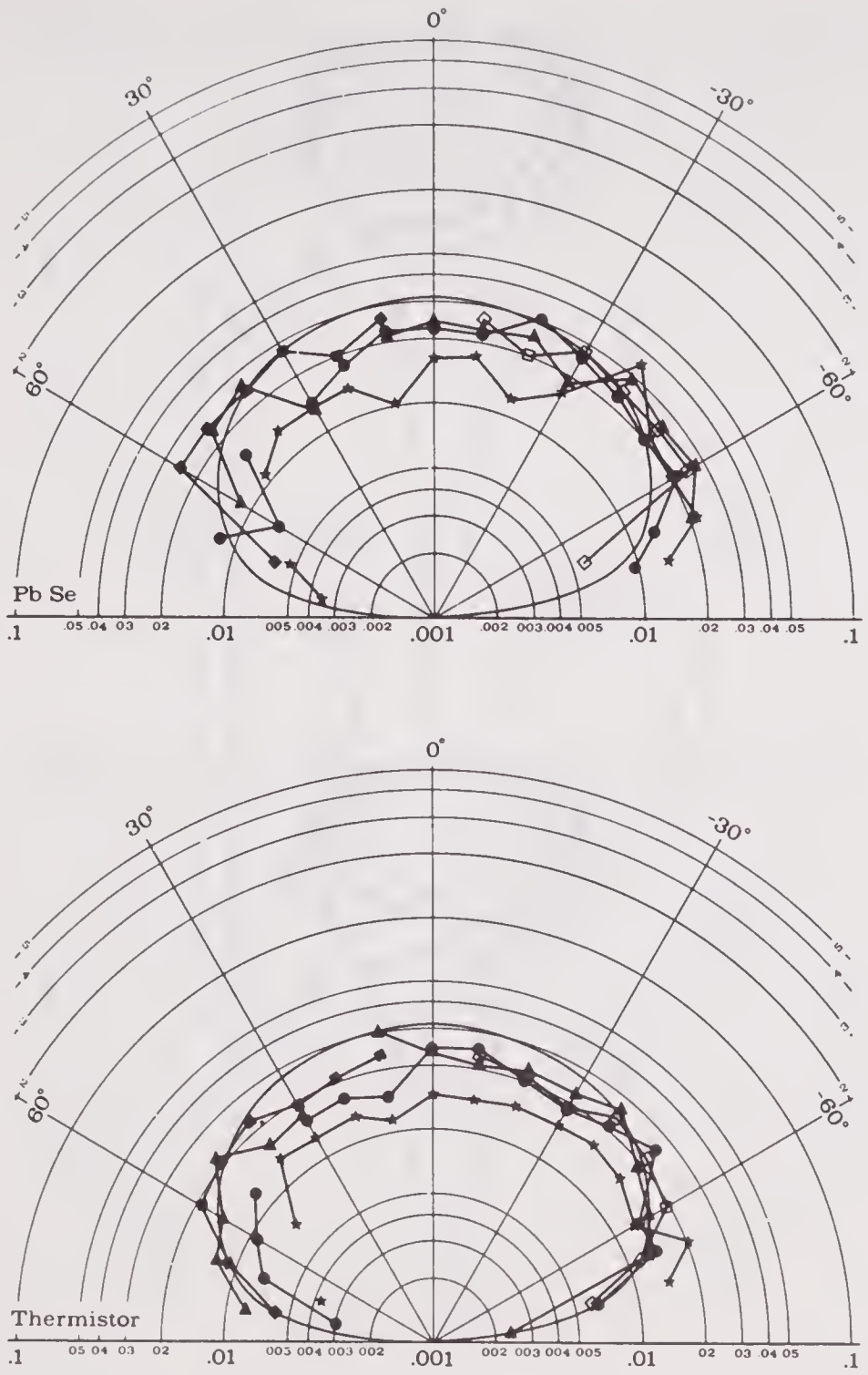


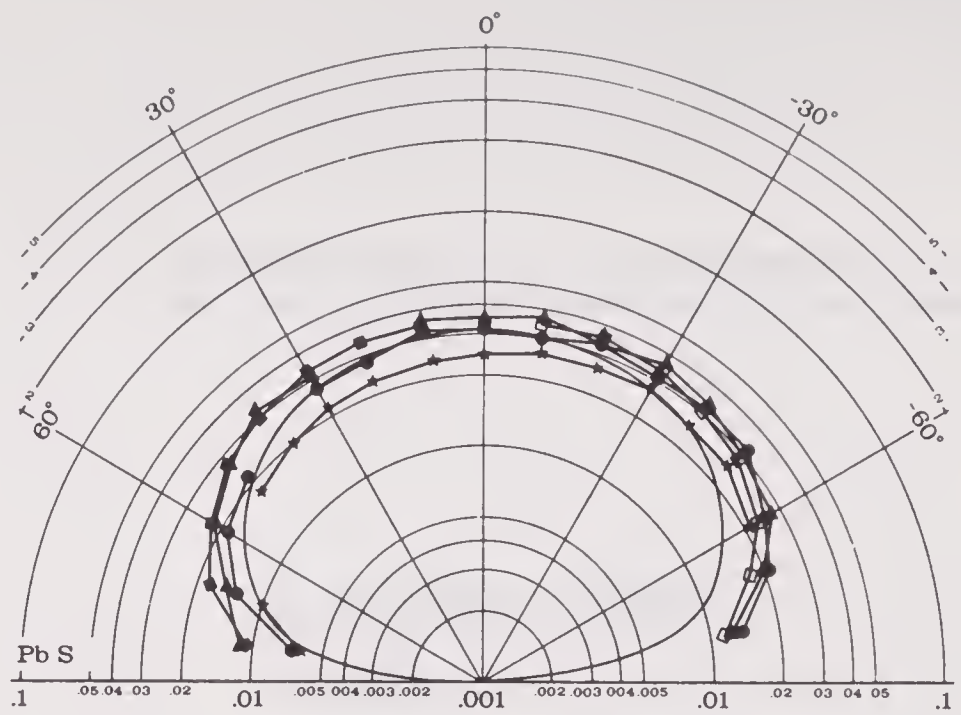
SOURCE ANGLES

- ■ 0°
- ▲ 20°
- ● 40°
- ★ 60°

Open Symbol - Theoretical
Closed Symbol - Experimental

FIG. 4-28. Directional reflectivity (Eq. 4-9) curves for concrete [17].

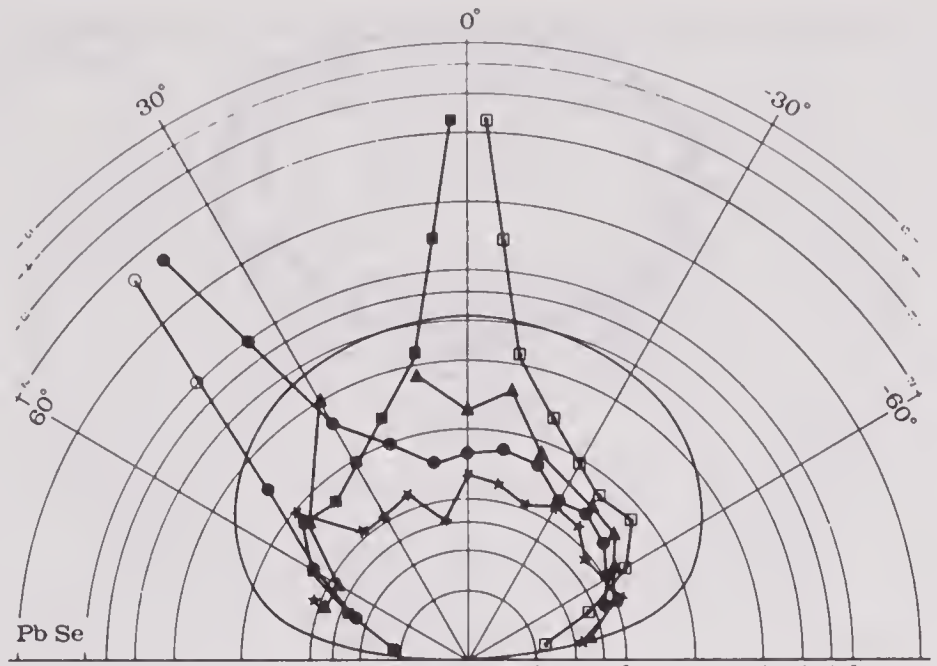


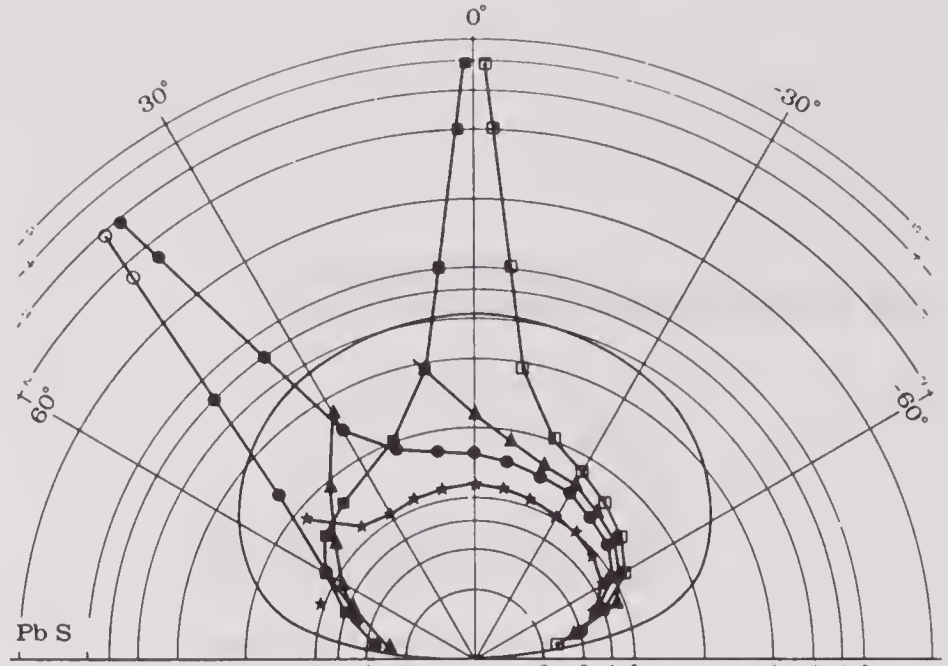


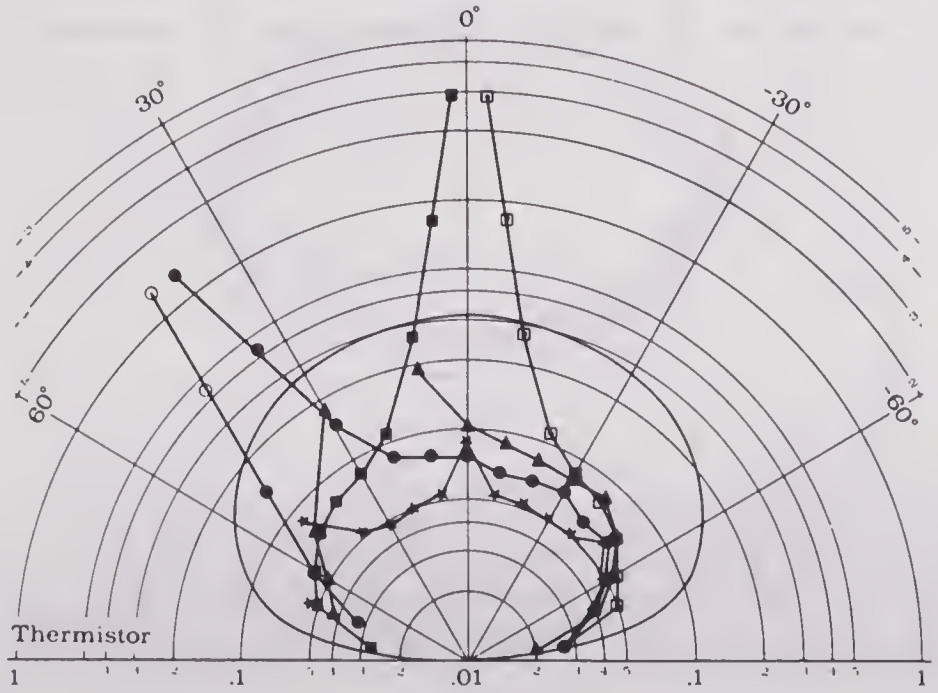
SOURCE ANGLES

■ 0°
▲ 20° Open Symbol - Theoretical
● 40° Closed Symbol - Experimental
★ 60°

FIG. 4-29. Directional reflectivity (Eq. 4-9) curves for asphalt [17].



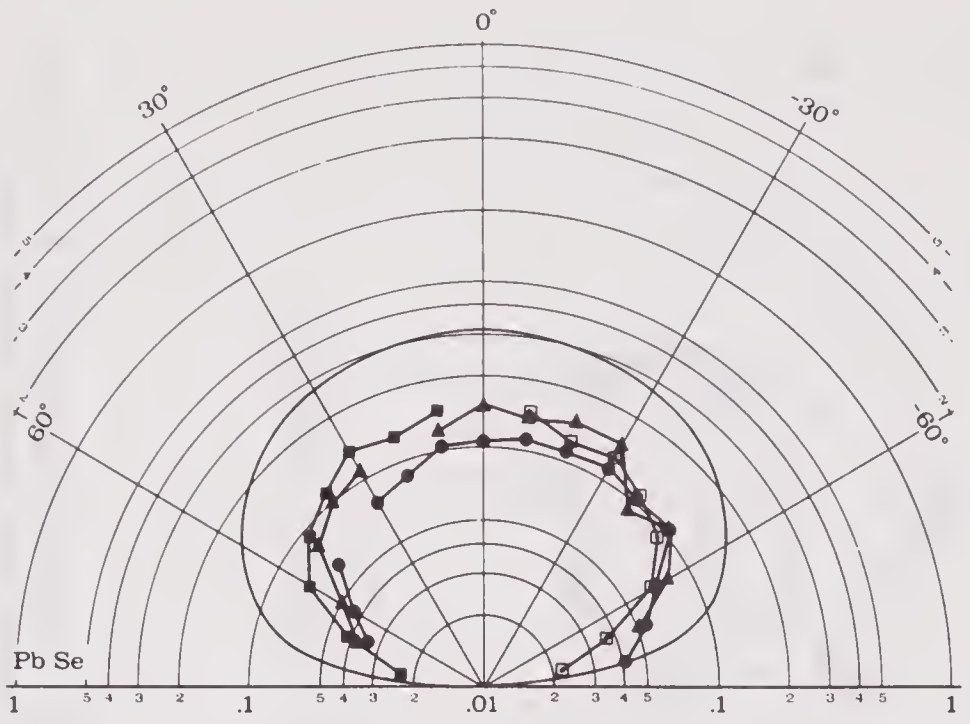


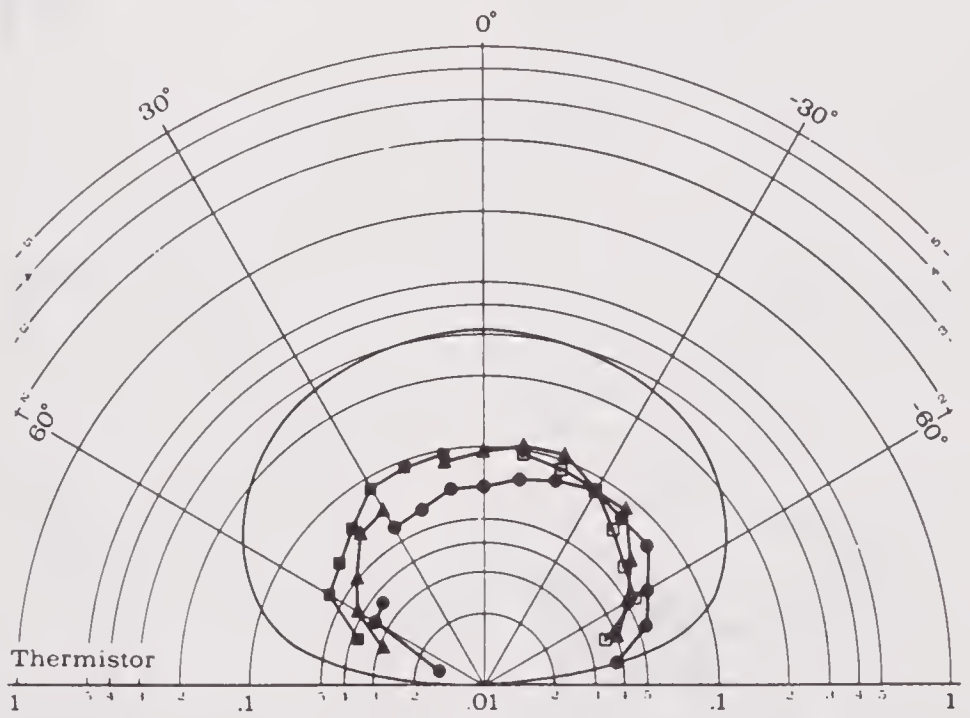


SOURCE ANGLES

- ■ 0°
- ▲ 20° Open Symbol - Theoretical
- ● 40° Closed Symbol - Experimental
- ★ 60°

FIG. 4-30. Directional reflectivity (Eq. 4-9) curves for concrete painted with Codit silver paint [17].



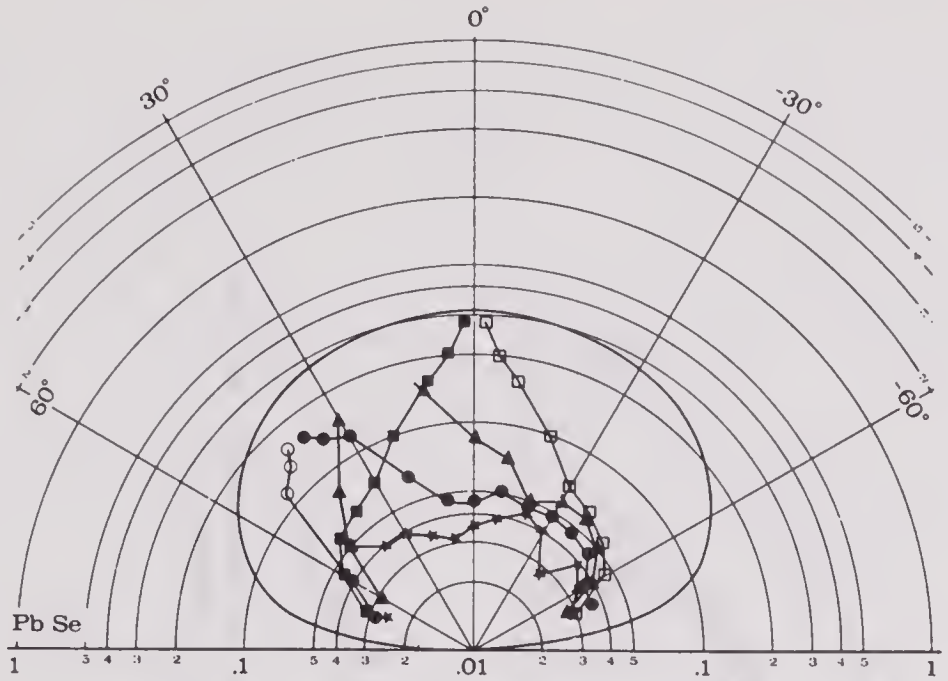


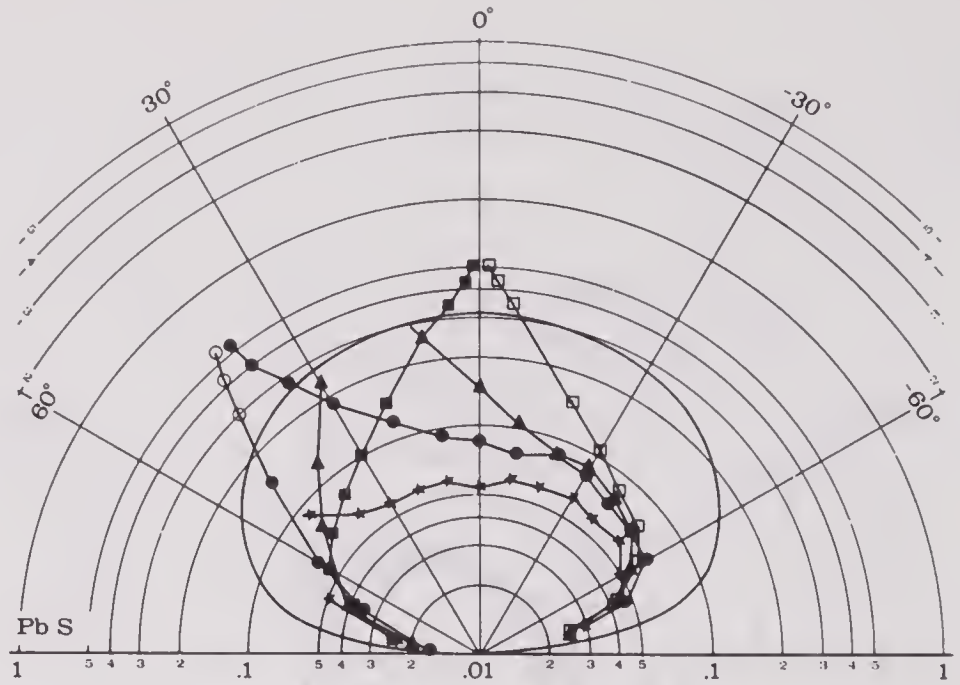
SOURCE ANGLES

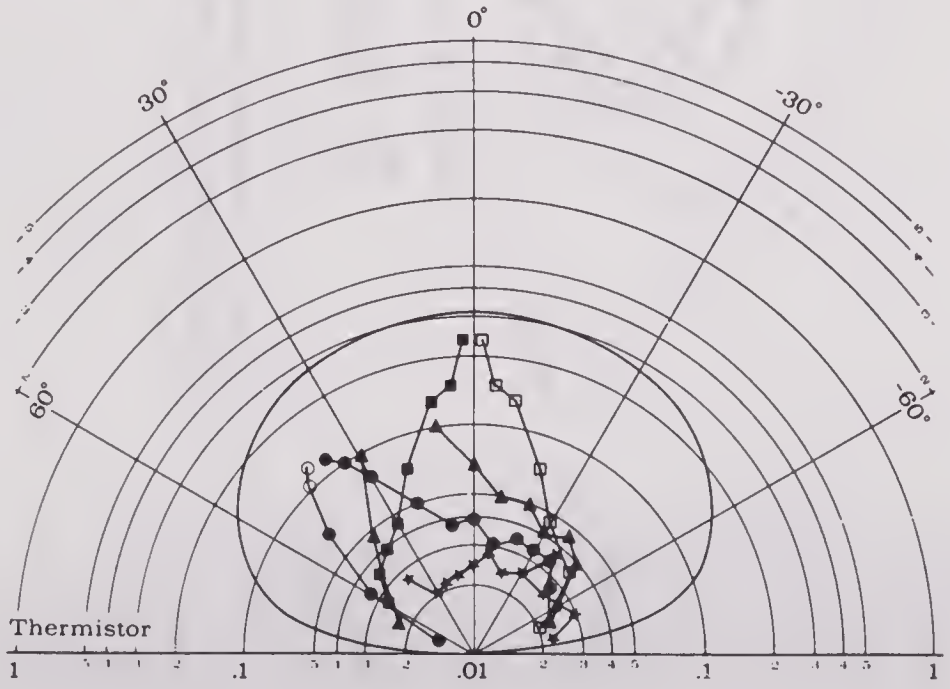
■ 0°
▲ 20°
● 40°
★ 60°

Open Symbol - Theoretical
Closed Symbol - Experimental

FIG. 4-31. Directional reflectivity (Eq. 4-9) curves for concrete painted with Centerlite white paint (abraded) [17].







SOURCE ANGLES
- ■ 0°
- ▲ 20°
- ● 40°
- ★ 60°

Open Symbol - Theoretical
Closed Symbol - Experimental

FIG. 4-32. Directional reflectivity (Eq. 4-9) curves for concrete painted with Flex-o-lite beaded paint [17].

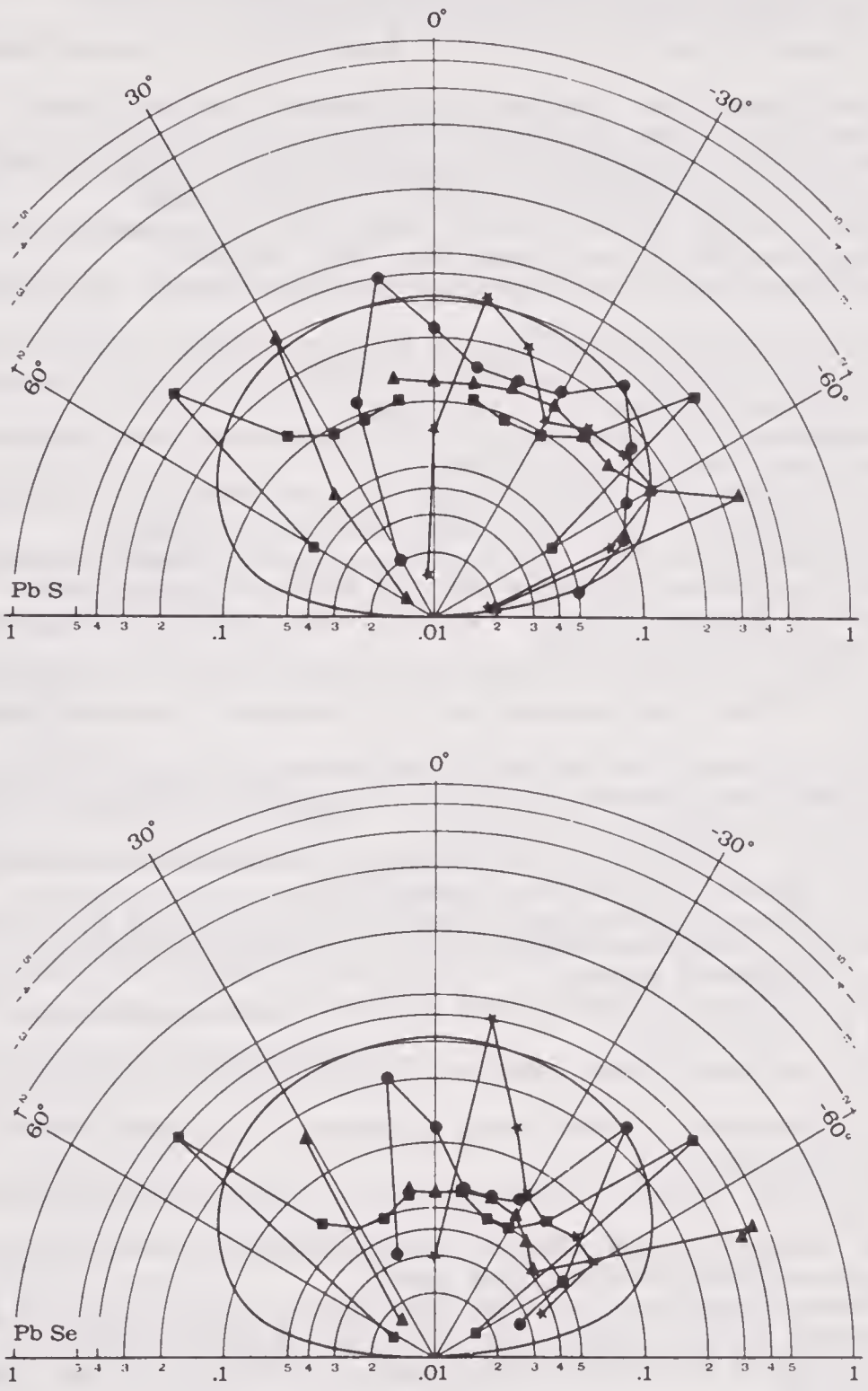


FIG. 4-33. Directional reflectivity (Eq. 4-9) curves for corrugated metal[17].

References

1. F. P. Boynton and J. T. Neu, *Rocket Plume Radiance,* V. *Calculation of Adiabatic Flame Temperatures of Afterburning Rocket Exhaust,* Rept. ERR-AN-011, Convair (Astronautics Division), San Diego, Calif. (May 1960).
2. W. H. Krase, *Exhaust-Gas Composition and Afterburning Energy Release for Selected Rocket Propellants,* Memo RM-3100-PR, Rand Corp. (April 1962).
3. *Design Information Manual* No. 101, Vol. I, Report SA57-53, Aerojet-General Corp. (Nov. 1, 1957).
4. *Rocket Exhaust Patterns,* Rept. No. 1844, Aerojet-General Corp. (Aug. 1960).
5. *Determination of the Envelope and Lines of Constant Mach Number for an Axially Symmetric Free Jet,* Rept. ZJ-7-054, Convair Division, General Dynamics (March 14, 1958).
6. G. P. Sutton, *Rocket Propulsion Elements,* Wiley, New York (1956).
7. Seth L. Tuttle, "A Model for the Radiation Field of the Atlas Sustainer at High Altitudes," *AMRAC Proceedings,* Vol. V, Part I, p. 69 (Nov. 1961) (SECRET).
8. G. Herzberg, *Infrared and Raman Spectra of Polyatonic Molecules,* Van Nostrand, New York (1945).
9. F. P. Boynton, *Rocket Plume Radiance,* IV. *Studies of Carbon Particles Formed by Small Hydrocarbon Fueled Rocket Engines,* Rept. ERR-AN-007, Convair (Astronautics Division), San Diego, Calif. (April 7, 1960).
10. D. D. Kurtovich and G. T. Pinson, *How to Find the Exhaust Heat Radiation of Aluminized Solid Rockets,* Space Aeronautics (July 1961).
11. *Transactions of BAMIRAC 1960 Summer Study,* Rept. No. 3768-6-X, Willow Run Laboratories, The University of Michigan, Ann Arbor, Mich. (Oct. 1960) (SECRET).
12. W. R. Fredrickson, N. Ginsburg, and R. Paulson, *Infrared Spectral Emissivity of Terrain,* WADC-TR-58-229, Syracuse University (April 30, 1958) AD 155 552.
13. N. Ginsburg, W. R. Fredrickson, and R. Paulson, "Measurements with a Spectral Radiometer," *J. Opt. Soc. Am.,* **50,** 1176 (Dec. 1960).
14. Dorothy E. Crowley, *Emittance and Reflectance in the Infrared; An Annotated Bibliography,* Rept. 2389-15-S, Willow Run Laboratories, The University of Michigan, Ann Arbor, Mich. (April 1959).
15. *Infrared Target and Background Radiometric Measurements—Concepts, Units, and Techniques* (Report of WGIRB), Rept. 2389-64-T, Institute of Science and Technology, The University of Michigan, Ann Arbor, Michigan (April 1959).
16. W. E. K. Middleton, "Vision Through the Atmosphere," in *Handbook der Physik,* ed. by Flugge, Vol. XLVIII, pp. 254 and 262, (Springer-Verlag (1957)).
17. A. LaRocca, J. Livisay, C. Miller, and G. Zissis, *Characteristics of a Ground Infrared Range,* Rept. No. 2849-6-F, Willow Run Laboratories, The University of Michigan, Ann Arbor, Mich. (June 1959) (SECRET) AD 314 492.
18. H. O. McMahon, "Thermal Radiation from Partially Transparent Reflecting Bodies," *J. Opt. Soc. Am.,* **40,** 376 (June 1950).
19. Henry H. Blau and Heinz Fischer, *Radiative Transfer from Solid Materials,* Macmillan, New York (1962).
20. Committee on Colorimetry, Optical Society of America, *The Science of Color,* Crowell, New York, pp. 176-178 (1954).
21. Air Force Cambridge Research Center, *Handbook of Geophysics for Air Force Designers,* Chap. 2 and 14.
22. G. T. Pullan, "Temperature and Emissivity Measurements," *Infrared Characteristics of RCN Destroyer-Escorts,* CARDE TR-413/62 (July 1962).
23. Geophysics Research Directorate, Air Force Cambridge Research Center, *U. S. Air Force Handbook of Geophysics,* revised ed., Macmillan, New York, pp. 215-220 (1960).
24. C. E. Heerema and H. C. Graboske, "Interpretation of Infrared Images," *Proc. IRIS,* **IV,** 4, 393-403 (Oct. 1959).
25. R. K. McDonald, "Techniques for Measuring Emissivities," *Proc. IRIS,* **V,** 3, 153 (July 1960).
26. H. H. Blau, Jr., J. L. Miles, and L. E. Ashman, *Thermal Radiation Characteristics of Solid Materials; A Review,* Scientific Report 1, AFCRC-TN-58-132, Arthur D. Little, Inc. (March 31, 1958). AD 146 883.
27. David K. Wilburn and Otto Renius, *The Spectral Reflectance of Ordnance Materials at Wavelengths of 1 to 12 Microns,* The Detroit Arsenal, Final Rept. No. 3196 (Feb. 1955). AD 087 246.
28. W. L. Starr, E. R. Streed, and A. I. Funai, *Principles of Infrared Camouflage for Low Temperature Targets,* Tech. Note N148, U. S. Naval Civil Engineering Research and Evaluation Laboratory, Port Hueneme, Calif. (July 21, 1953).
29. Max Kronstein, *Research Studies and Investigations on Spectral Reflectance and Absorption Characteristics of Camouflage Paint Materials and Natural Objects,* Final Report, New York University (March 1955). AD 100 058.
30. G. D. Currie *et al., Infrared Aerial Reconnaissance Interpretation,* Bendix Corp., Systems Division, Ann Arbor, Mich., Note BSR-175, RADC-TN-60-29 (January 24, 1960) (CONFIDENTIAL). AD 315 971.

Chapter 5

BACKGROUNDS

Richard Kauth

The University of Michigan

CONTENTS

5.1. Sky Backgrounds 96
5.2. Aurora 100
5.2.1. Auroral Spectra 100
5.2.2. Auroral Zones 101
5.2.3. Periodic Variations 102
5.2.4. Height and Vertical Extent 102
5.3. Night Airglow 104
5.4. Stellar Radiation 107
5.4.1. Stellar Magnitudes 107
5.4.2. Stellar Spectral Classes 107
5.4.3. Number of Stars 107
5.4.4. Galactic Concentration of Stars 109
5.4.5. Spectral Distribution of Stellar Radiation 110
5.4.6. Determining Spectral Irradiance of Celestial Bodies 112
5.5. The Earth as a Background 115
5.5.1. Geometric Relationships 115
5.5.2. Path Lengths 116
5.6. Cloud Meteorology 118
5.6.1. Cirrus Clouds 122
5.6.2. Stratospheric Clouds 124
5.6.3. Probability of Coverage at Various Altitudes 124
5.7. Stratospheric Aerosols 141
5.8. Spectral Radiance of Terrain 142
5.8.1. Terrain Temperature 142
5.8.2. Terrain Emissivity and Reflectivity 142
5.8.3. Spectra in the Emission Region 143
5.8.4. Spectra in the Scattering Region 145
5.8.5. Spectral Radiance of Various Objects and Surfaces 145
5.9. Marine Backgrounds 166
5.9.1. Infrared Optical Properties of Sea Water 166
5.9.2. Sea-Surface Geometry 168
5.9.3. Sea-Surface Temperature Distribution 169
5.9.4. Sky Radiance 170
Notes Added in Proof 170

5. Backgrounds

5.1. Sky Backgrounds*

Sky-background radiation in the infrared is caused by scattering of the sun's radiation and by emission from atmospheric constituents. Figure 5-1 illustrates the separation of the spectrum into two regions—the solar scattering region short of 3 μ, and the thermal emission region beyond 4 μ. Solar scattering is represented by reflection from a bright sunlit cloud, and alternatively by a curve for clear-air scattering. The thermal region is represented by a 300°K blackbody. Figure 5-2 shows blackbody curves for temperatures ranging from 0° to 40°C. This simple model is modified by a number of factors: in the solar region there are absorption bands of water vapor at 0.94, 1.1, 1.4, 1.9, and 2.7 μ, and of carbon dioxide at 2.7 μ. The effect of these bands is shown in Fig. 5-3.

In the thermal region the bands which have strong absorption, and thus strong emission, will approach very closely to the blackbody curve appropriate to the temperature of the atmosphere. Less strongly emitting regions may contribute only a small fraction of the radiation of a blackbody at the temperature of the atmosphere. The bottom curve in Fig. 5-4 is a good example. This zenith measurement, taken from a high, dry location, Elk Park, Colorado, shows low emission except in the strong band of CO_2 at 15 μ and of H_2O at 6.3 μ. There is also a weak emission peak, due to ozone,

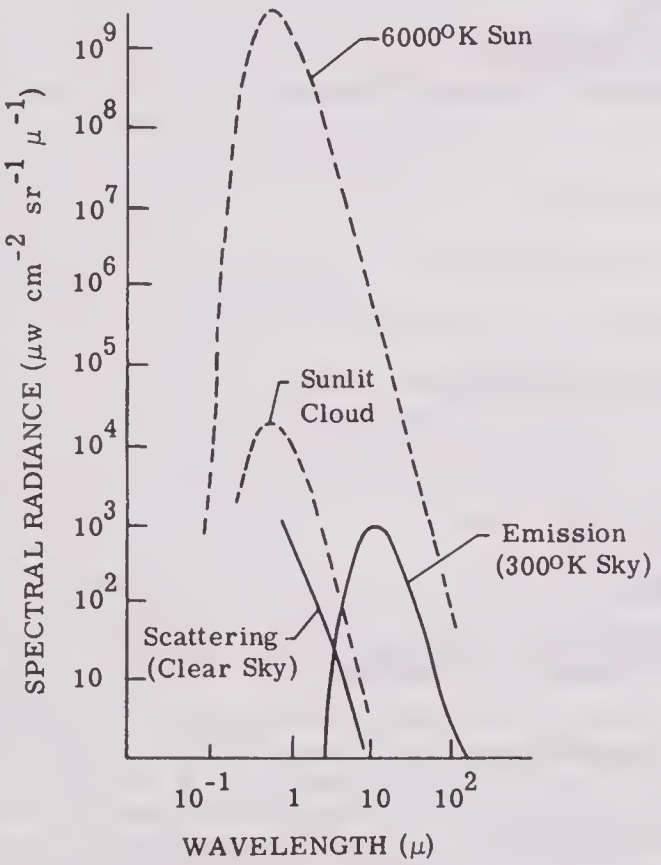


FIG. 5-1. Contributions from scattering and atmospheric emission to background radiation [1].

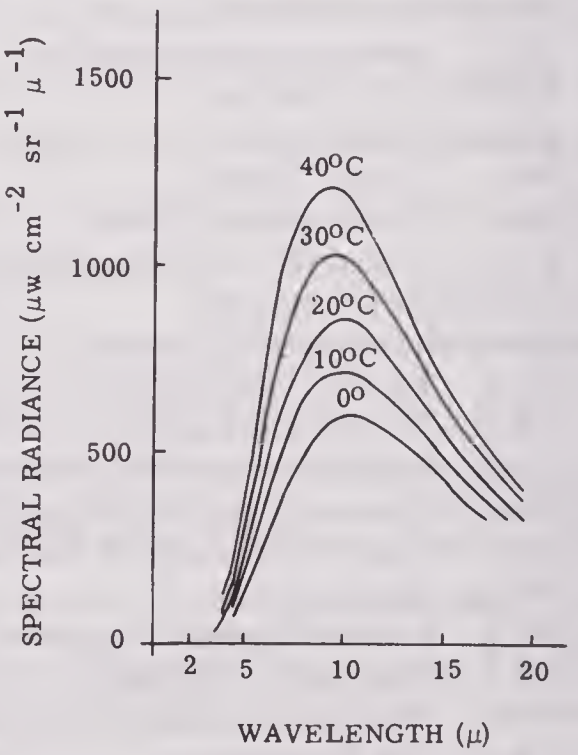


FIG. 5-2. Spectral radiance of a blackbody with temperature in the range of 0 to 40°C [1].

*See also *Notes Added in Proof*, page 170.

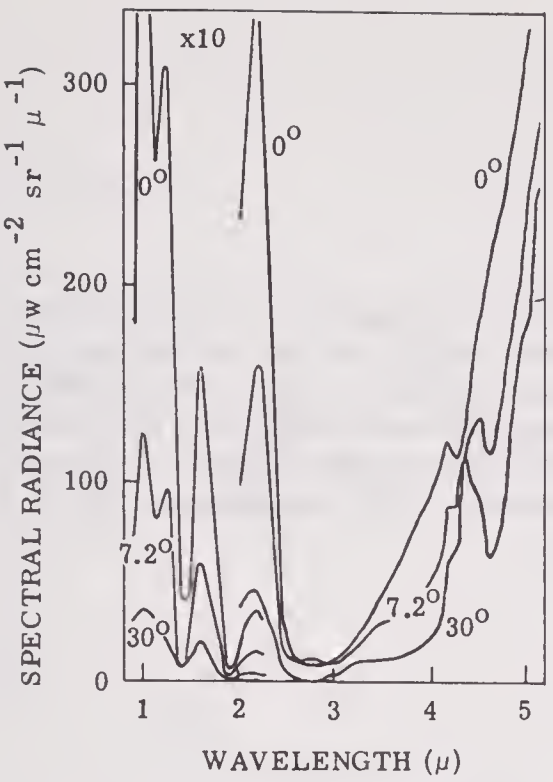


FIG. 5-3. Spectral radiance of the clear daytime sky [1].

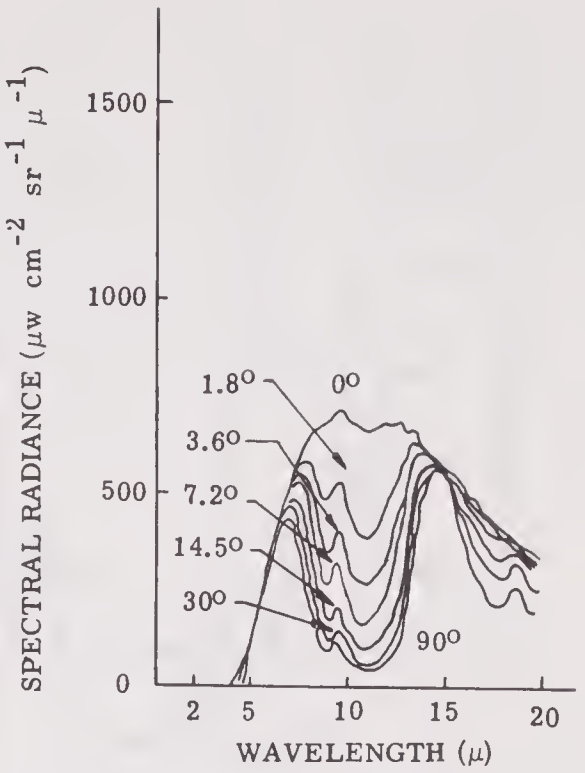


FIG. 5-4. The spectral radiance of a clear nighttime sky for several angles of elevation above the horizon (Elk Park Station, Colorado) [1].

at 9.6 μ. The low-level continuum is due to the wings of the strong bands of H_2O and CO_2. The effect of increased humidity and air mass can be seen by comparing the bottom curves of Fig. 5-4 and 5-5. Figure 5-5 shows measurements taken at a humid sea-level location, Cocoa Beach, Florida.

The effect of increasing air mass alone can be seen in both Fig. 5-4 and 5-5, by comparing curves taken from the same altitude at various elevation angles. The emission shows a systematic decrease with increasing elevation angle. The direction of look also has an effect in the solar scattering region, as seen in Fig. 5-3, where, for a clear sky, the sun position is fixed and the spectral radiance is plotted for several observer angles.

The position of the sun has a strong effect on the scattered radiation in the solar region, as shown in Fig. 5-6, where the observer looks at the zenith and the elevation angle of the sun is varied but has little effect on the radiation in the thermal region. The temperature of the atmosphere, on the other hand, has a strong effect on the radiation in the thermal region but little effect in the solar region. The presence of clouds will affect both the near-infrared solar scattering and the thermal-region emission.

Near-infrared radiation exhibits strong forward scattering in clouds. Thus the relative positions of sun, observer, and cloud become especially important. For a heavy overcast sky, multiple scattering reduces the strong forward scattering effect.

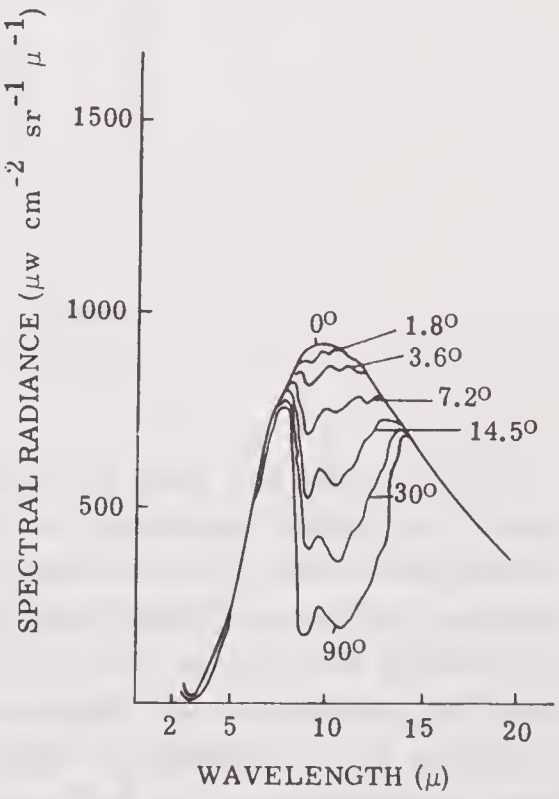


FIG. 5-5. The spectral radiance of a clear nighttime sky for several angles of elevation above the horizon (Cocoa Beach, Florida) [1].

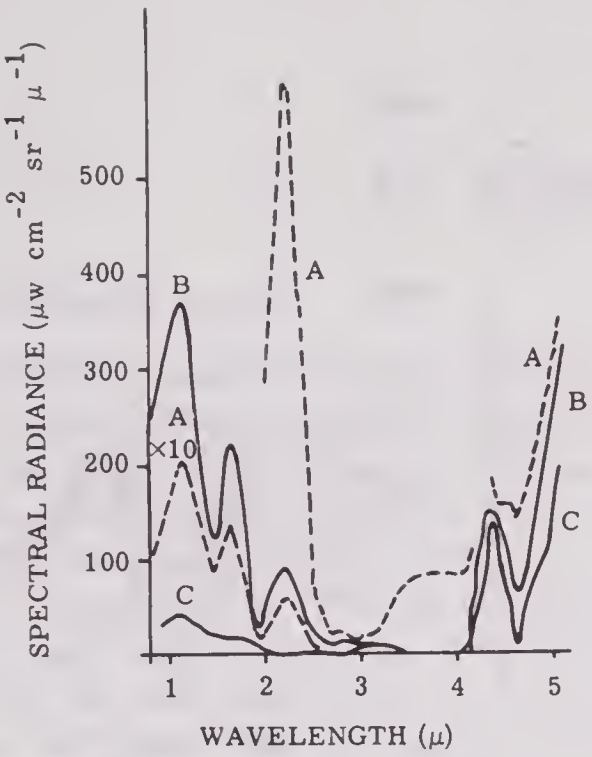


FIG. 5-6. Spectral radiance of a clear zenith sky as a function of sun position; *A* = sun elevation 77°, temperature 30°C; *B* = sun elevation 41°, temperature 25.5°C; *C* = sun elevation 15°, temperature 26.5°C [1].

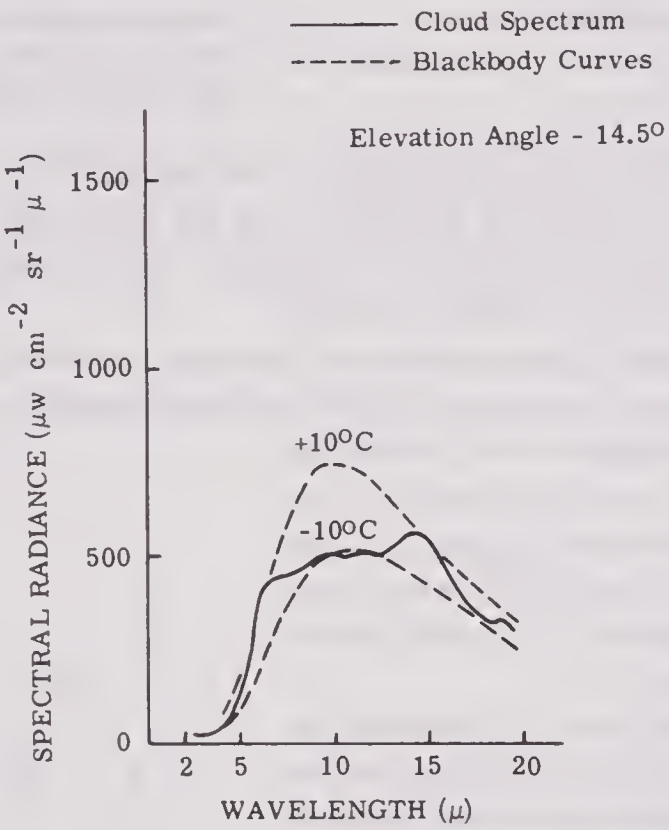


FIG. 5-7. The spectral radiance of the underside of a dark cumulus cloud [1].

Thick clouds are good blackbodies. Emission from clouds is in the 8-13-μ region and is, of course, dependent on the cloud temperature. Because of the emission and absorption bands of the atmosphere at 6.3 μ and 15 μ, a cloud may not be visible in these regions and the radiation here is determined by the temperature of the atmosphere. A striking example is given in Fig. 5-7. Here the atmospheric temperature is +10°C and the radiation in the emission bands at 6.3 μ and 15 μ approaches a value appropriate to that temperature. The underside of the cloud has a temperature of −10°C, and the radiation in the 8-13-μ window approaches that of a blackbody at −10°C.

Figure 5-8 shows the variation of sky radiance as a function of elevation angle. Figure 5-9 shows the variation with respect to variations of ambient air temperature, and Fig. 5-10 shows seasonal variations.

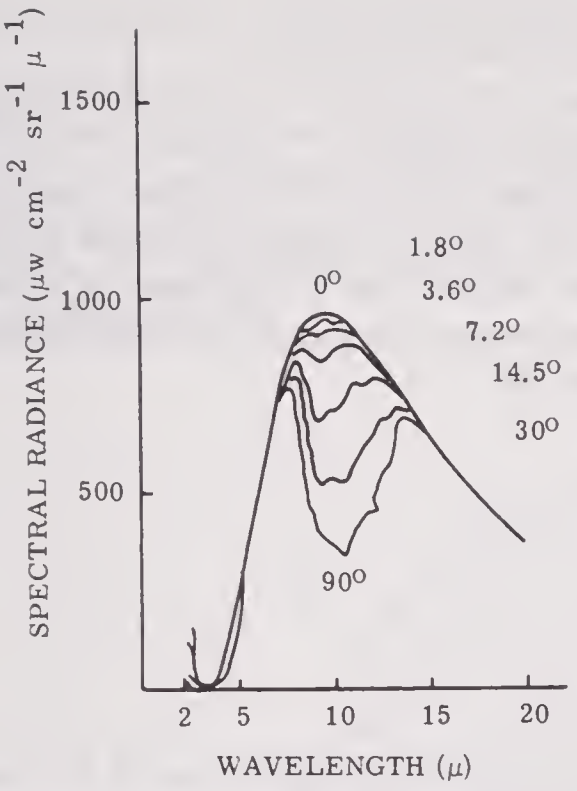


FIG. 5-8. The spectral radiance of sky covered with cirrus clouds at several angles of elevation [1].

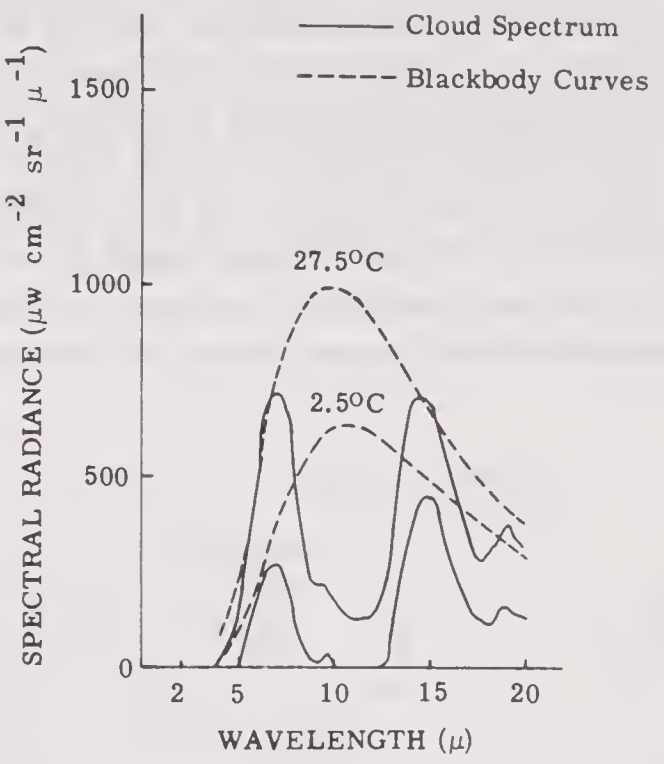


FIG. 5-9. Zenith sky spectral radiance showing the large variation with ambient air temperature [1].

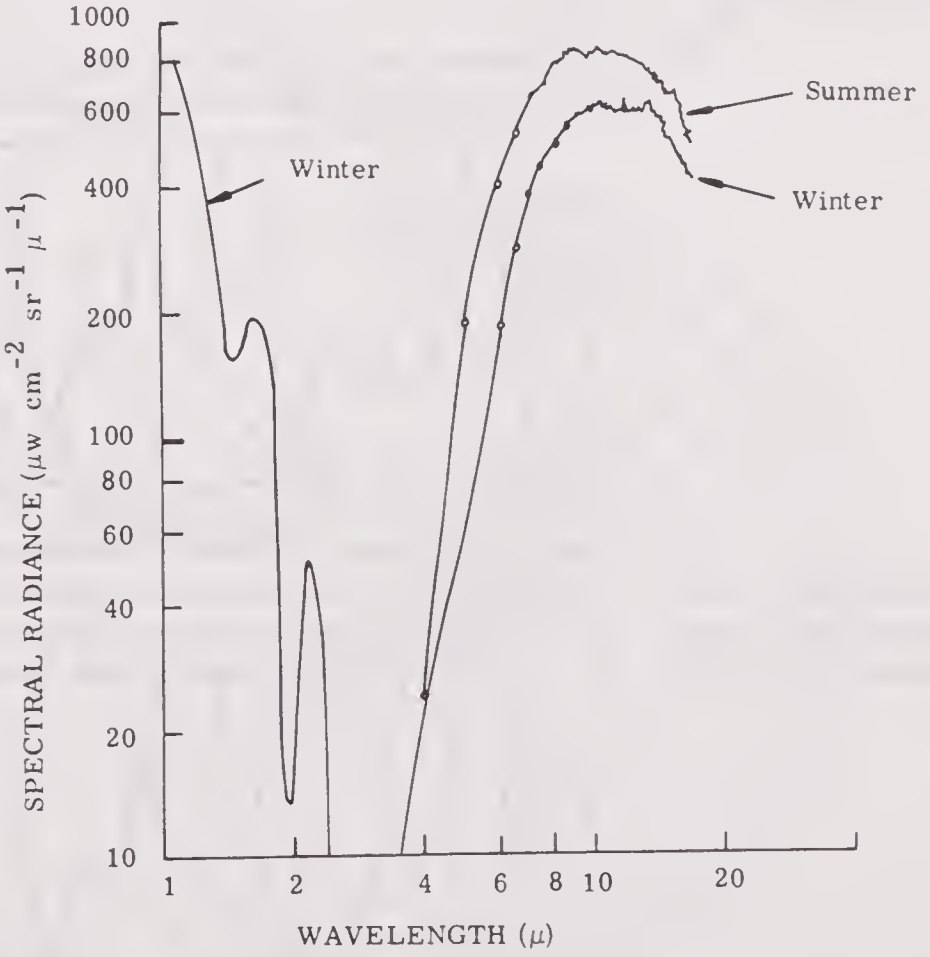


FIG. 5-10. Spectral radiance of overcast skies in winter and summer [2].

5.2. Aurora [3,4].

5.2.1. Auroral Spectra.* Aurora emission lines occur at 0.92, 1.04, and 1.11 μ; the measured brightnesses are about 6×10^{-8} w cm^{-2} sr^{-1} line^{-1} [5].

Figure 5-11 shows the auroral spectrum between 0.9 and 1.2 μ. This reproduction was obtained by averaging a number of individual spectra [5]. Figure 5-12 shows auroral spectra between 1.4 and 1.65 μ. The dotted curve is the airglow spectrum fitted to the auroral spectrum in a region where the auroral emission appears feeble. Spectra (*a*), (*b*), and (*c*) were made in consecutive scans, with a total time of 3 min. The relative intensities of features on a single scan are not significant, since the aurora fluctuates in brightness during the scanning period [6].

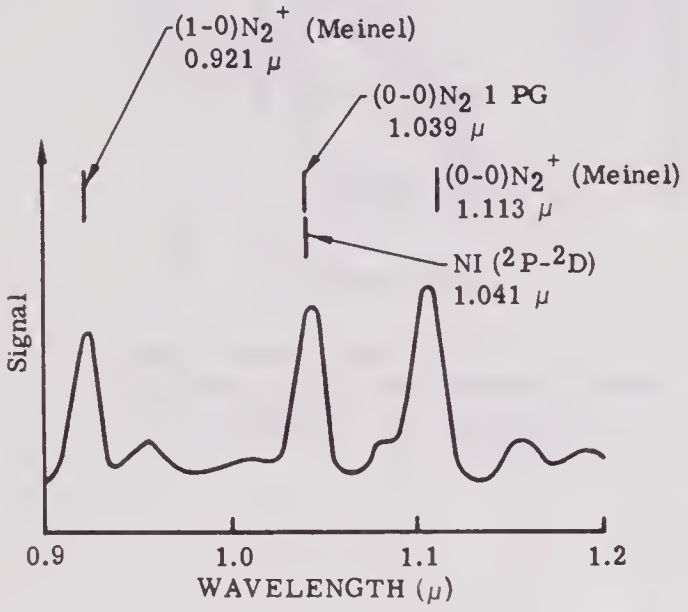
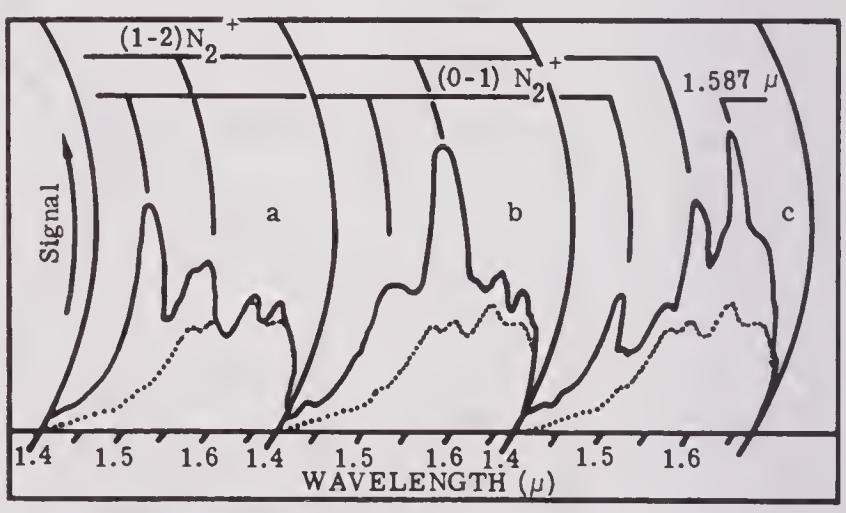


FIG. 5-11. Auroral spectrum, 0.9 to 1.2 μ, obtained with a lead sulfide spectrometer; projected slit width, 100 A [5].



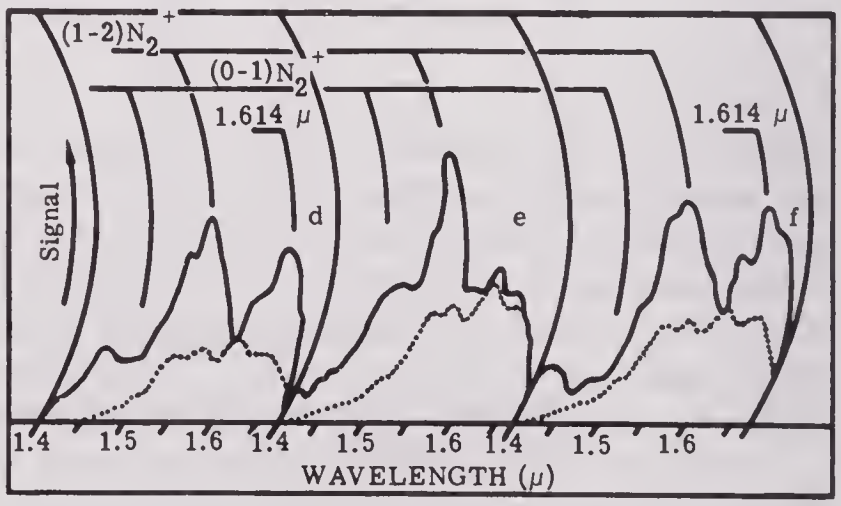


FIG. 5-12. Auroral spectra, 1.4 to 1.65 μ, obtained with a lead sulfide spectrometer; projected slit width 200 A [6].

*See also *Notes Added in Proof*, page 170.

5.2.2. Auroral Zones [4]. Figure 5-13 shows the auroral zones. These are divided into 3 areas: the north and south auroral regions extending from geomagnetic latitudes 60° to the poles, the subauroral belts between 45° and 60°, and the minauroral belt between 45°N and 45°S.

The auroral regions include the auroral zones, which are the regions of maximum occurrence, and the auroral caps, which are the polar regions within the auroral zones.

Although aurorae occur primarily in the auroral regions, large displays may occur in quite low latitudes. However, in tropical and even low temperate latitudes they are extremely rare.

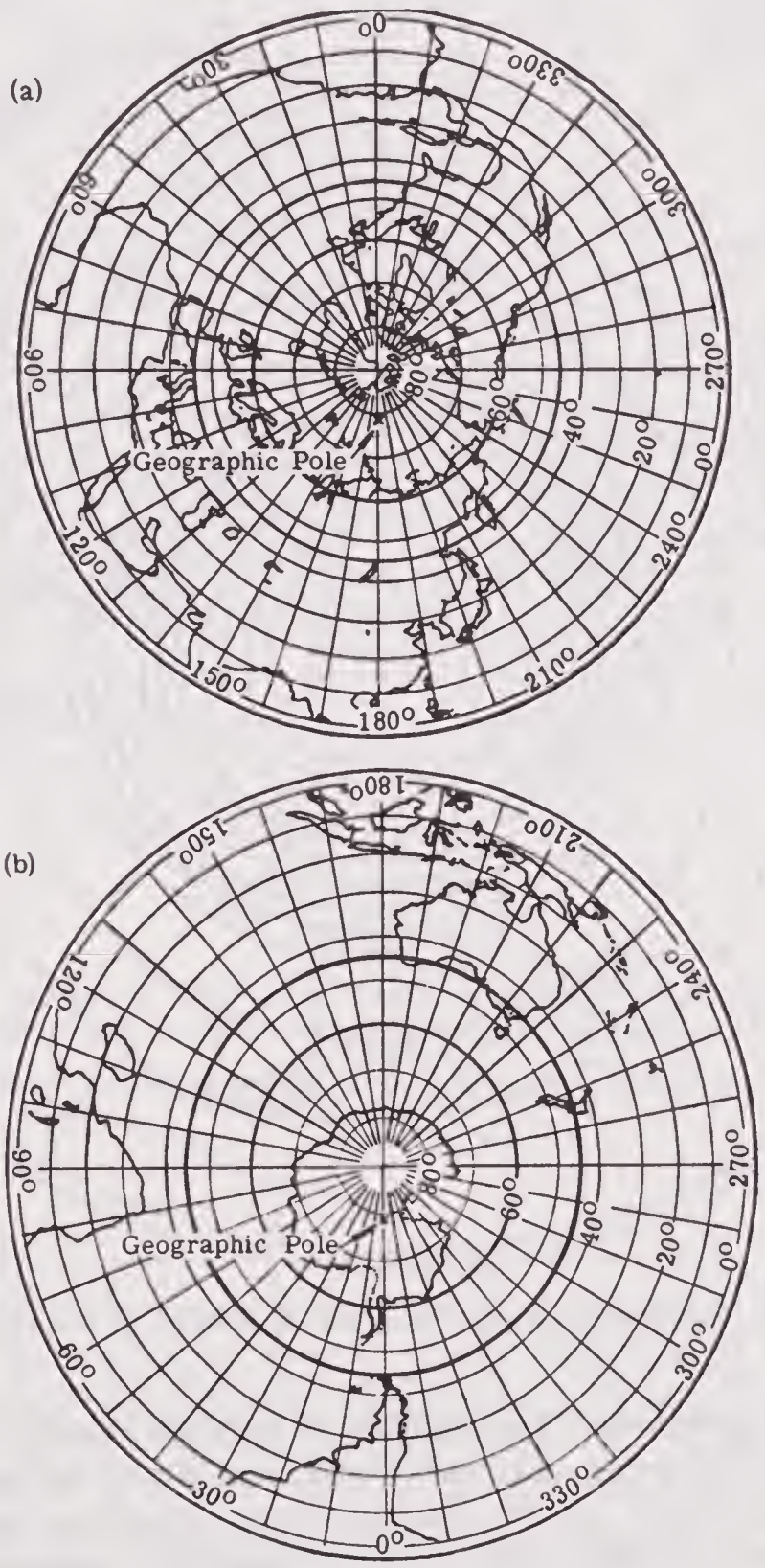


FIG. 5-13. The hemisphere centered on (*a*) northern geomagnetic pole (78°5′N, 69°W geographic) and (*b*) southern geomagnetic pole (78°5′S, 111°E geographic) [4].

The frequency of auroral occurrences has a maximum some 20° or 25° from the geomagnetic poles. Figure 5-14 shows the geographic distribution of the frequency of aurorae in the northern hemisphere [7]. The isochasms refer to the number of nights during the year in which an aurora might be seen at some time during the night, and in any part of the sky, if clouds and other factors affecting visual detection of aurorae do not interfere. Figure 5-15 shows the zone of maximum auroral frequency in the southern hemisphere [8].



FIG. 5-14. Geographic distribution of the frequency of aurorae in the northern hemisphere [7].

5.2.3. Periodic Variations. The number of aurorae observed from a particular point over the course of a year may vary widely and is strongly correlated with solar activity. Minimum auroral activity corresponds with minimum solar activity. Maximum auroral activity usually occurs about two years after sunspot maximum.

5.2.4. Height and Vertical Extent. On auroral arcs and bands the most convenient height to measure is the apparent lower border, which is fairly sharp. An example of a set of such measurements in and near the auroral zone is shown in Fig. 5-16 [9].

The total number of measurements shows a concentration between 95 and 110 km, with a double peak. The lower limits of individual rays appear 10 or 15 km higher than the lower edges of most arcs, bands, and draperies. Sunlit auroral rays appear systematically higher than displays in the dark atmosphere. Figure 5-17 shows the heights of rays over southern Norway. A few sunlit rays extend higher than 1000 km.

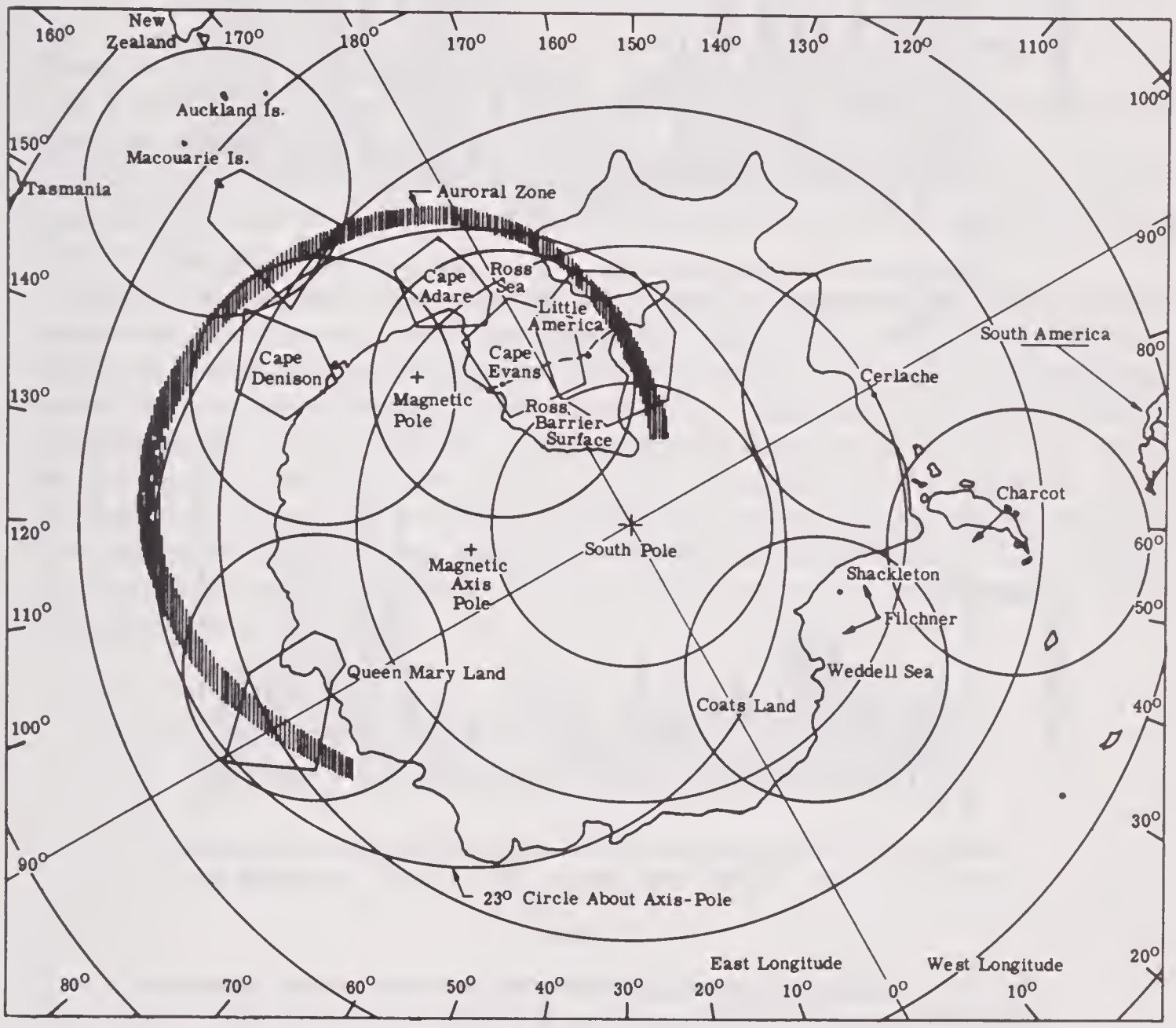


FIG. 5-15. Zone of maximum auroral frequency in the southern hemisphere [8].

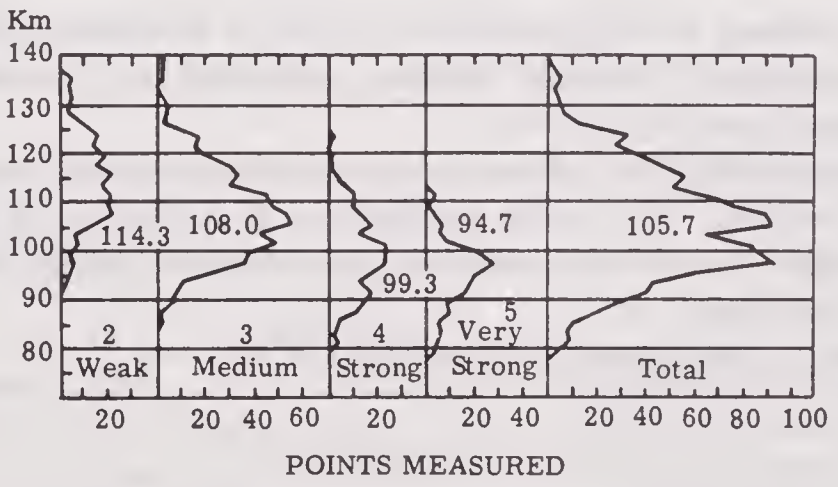


FIG. 5-16. Distribution of heights of lower borders of auroral arcs [9].

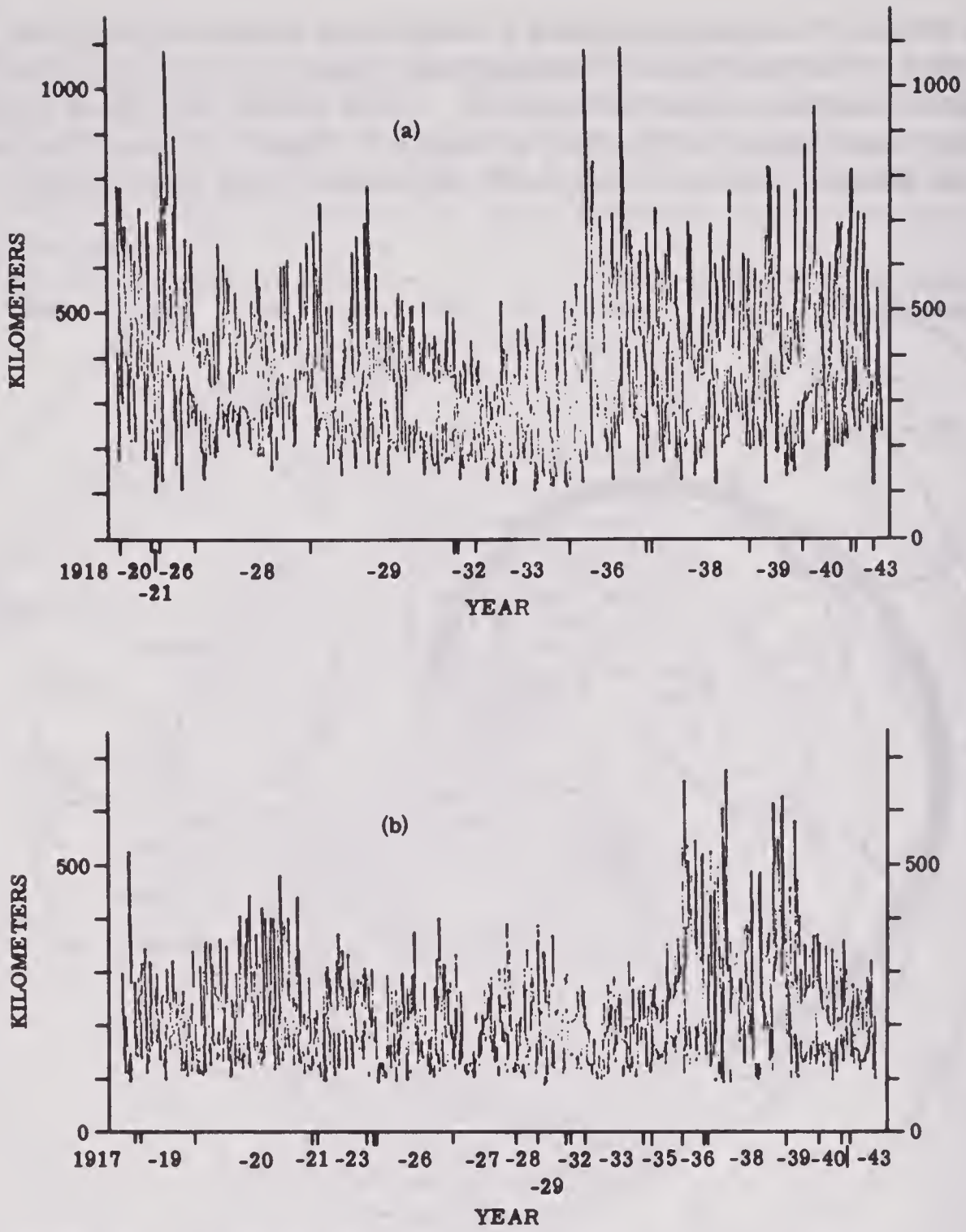


FIG. 5-17. Length and position in the atmosphere of the vertical projections of auroral rays in (*a*) sunlight and (*b*) earth's shadow (1917-1943) [10].

5.3. Night Airglow

Airglow may be defined as the nonthermal radiation emitted by the earth's atmosphere, with the exceptions of auroral emission and radiation of a cataclysmic origin, such as lightning and meteor trails [4].

Night airglow emissions in the infrared are caused by transitions between vibrational states of the OH^- radical. The exact mechanism of excitation is still unclear, but the effect is to release energy from solar radiation stored during the daytime. Airglow occurs at all latitudes.

There is evidence [11] that some of the excitation is

$$H + O_3 \rightarrow OH + O_2$$

$$OH + O \rightarrow O_2 + H$$

Thus, it appears that the distribution of night airglow is related to that of ozone. The measured heights of the airglow range from 70 to 90 km, which corresponds to the location of ozone.

Airglow brightness is specified in rayleighs (R) a measure of the apparent number of photons emitted in a column 1 cm^2 in diameter along the observer's line of sight.

$$1\ \mathrm{R} = 10^6\ \text{(apparent) photon/cm}^2\text{-sec (column)} = 4\pi I$$

where I is millions of photon/(cm^2-sec-sr).

To a good approximation, the nightglow increases away from the zenith as sec θ. Measurements usually are reported normalized to the zenith.

Variations in airglow intensity during the night seem to be caused by the motion of large patches (airglow "cells") with dimensions of about 2500 km moving with velocities of about 70 m sec^{-1} [12].

Figure 5-18 shows the relative brightness of airglow intensity [13]. Airglow emissions due to OH^- appear as small maxima in the vicinity of 1.6 μ and 2.15 μ. Although further emission bands are predicted in the range from 2.8 to 4.5 μ, they are thoroughly masked by the thermal emission of the atmosphere. Figure 5-19 shows the nightglow spectrum in the 1-2-μ region [14]. Looking straight down from a satellite, the atmospheric spectrum should be very similar to that shown in Fig. 5-18 and 5-19. Table 5-1 compares the approximate rates of emission for various airglow and auroral lines. The references in the footnotes should be consulted for further details. Note that, for the airglow, all results are given for the zenith itself rather than for the angles at which observations are usually made.

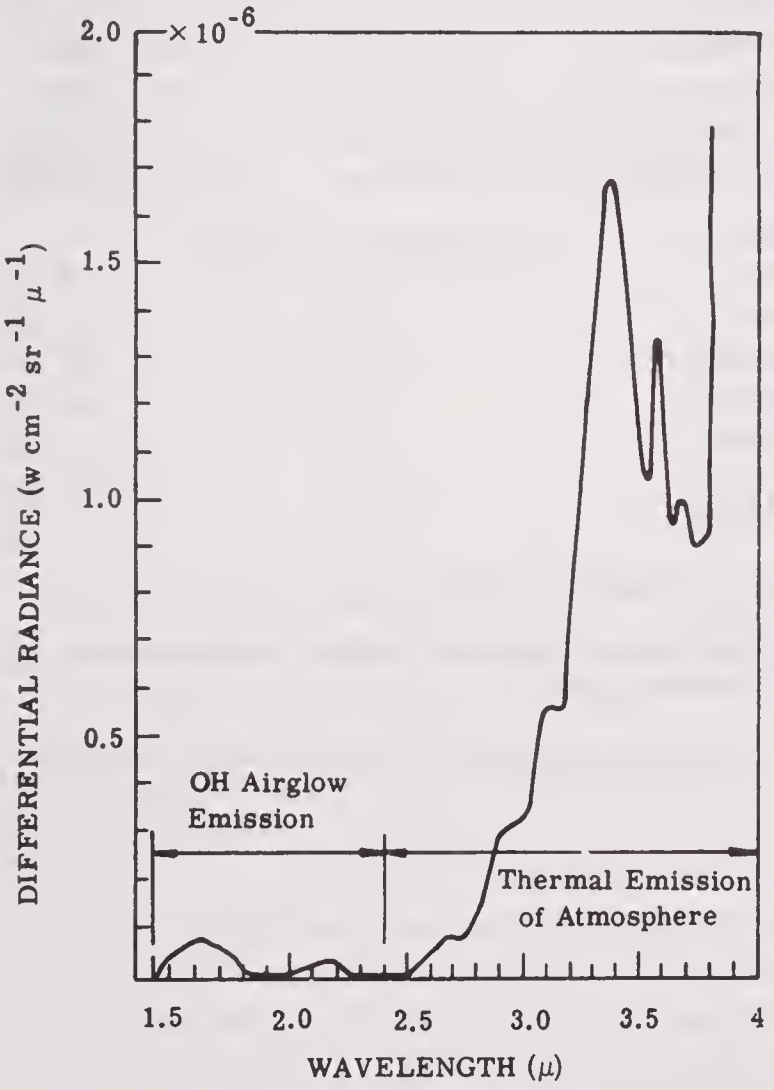


FIG. 5-18. Airglow intensity [13].

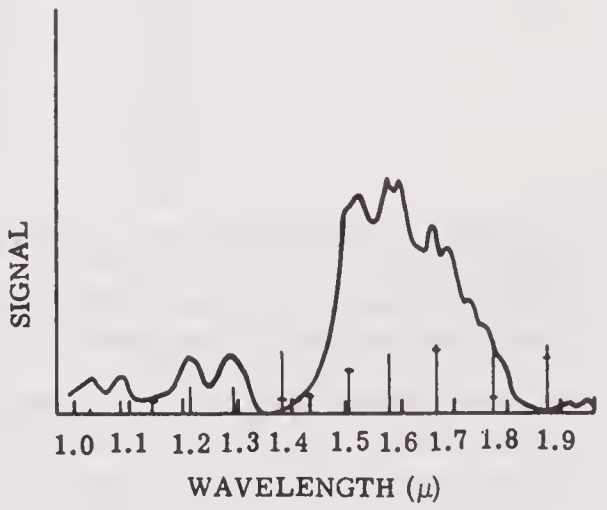


FIG. 5-19. Nightglow spectrum, obtained with a scanning spectrometer (projected slit width 200 A). The origins and expected intensities of OH bands are shown by vertical lines; the horizontal strokes indicate the reduction due to water vapor [14].

TABLE 5-1. COMPARISON OF AURORA AND AIRGLOW PHOTON EMISSION RATES [4]

Source	*Emission*	$4\pi I^e$
Aurora,[a] IBC I	$[OI]_{32}$ 5577 A	1 kR
II	–	10 kR
III	–	100 kR
IV	–	1000 kR
Night airglow[b] (in the zenith)	$[OI]_{32}$ 5577 A	250 R
	$[OI]_{21}$ 6300 A	50-100 R
	Na 5893 A	–
	summer	< 30 R
	winter	200 R
	Hα 6563 A	5-20 R
	Lγ α 1215 A	2.5 kR
	O_2 Atmospheric (0-1) 8645 A	1.5 kR
	O_2 Herzberg (observable range)	430 R
	OH (4-2) 1.58 μ	175 kR
	OH (estimated total)	4500 kR
Twilight airglow[c] (referred to the zenith)	N_2^+ 3914 A	1 kR
	(quiet magnetic conditions)	–
	Na I 5893 A	–
	summer	1 kR
	winter	5 kR
	$[OI]_{21}$ 6300 A	1 kR
	Ca II 3933 A	150 R
	Li I 6708 A	200 R
	$[NI]_{21}$ 5199 A	10 R
	O_2 IR Atmospheric (0-1) 1.58 μ	20 kR
Day airglow[d] (referred to the zenith)	Na 5893 A	–
	summer	2 kR
	winter	15 kR
	$[OI]_{21}$ 6300 A	50 kR
	OI 8446 A	0.5 kR
	OI 11,290 A	0.5 kR
	N_2^+ 3914 A	< 70 kR > 1 kR

[a] Recommended as definitions of the International Brightness Coefficients (IBC) [15,16].
[b] Average values.
[c] Approximate values of the maximum emission rates that are observed during twilight. These values are often governed by the time after sunset when observations first become possible.
[d] Values predicted from theory [17-20].
[e] $4\pi I$ is the apparent emission rate in rayleighs.1 R = an apparent emission rate of 1 megaphoton/cm²-sec (column).

Figure 5-20 shows the frequency distribution of air glow and weak auroral brightnesses near the geomagnetic pole (Thule, Greenland) and at a subauroral station (Fritz Peak, Colorado [9].

There is some evidence that suggests a general increasing brightness of airglow emissions toward higher latitudes and a bright belt at middle latitudes.

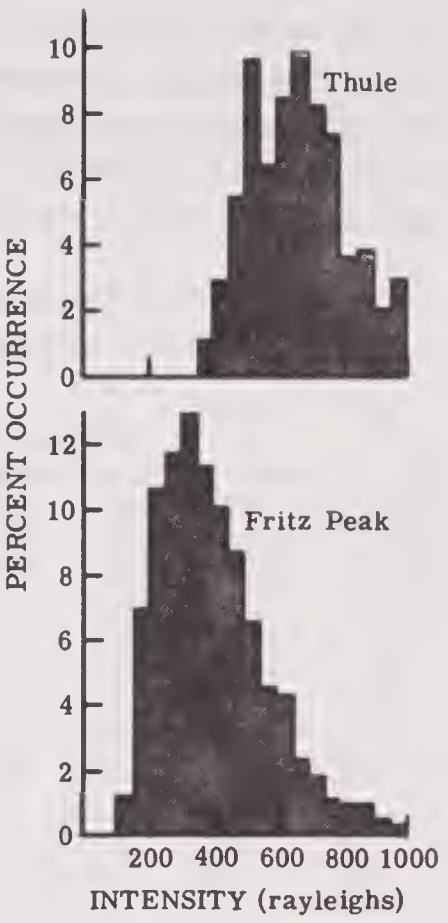


FIG. 5-20. Frequency distribution of airglow [9].

5.4. Stellar Radiation

5.4.1. Stellar Magnitudes [21]. The brightness of celestial bodies is usually measured in magnitudes. The scale of magnitudes is adjusted so that a star of magnitude +1.00 (first magnitude) gives a luminous flux of 0.832×10^{-10} lumen cm^{-2} at a point outside the atmosphere of the earth.

The relation between the visible light received from two stars and their magnitudes is expressed by the formula

$$\log \frac{I_1}{I_2} = 0.4(m_2 - m_1) \qquad (5\text{-}1)$$

where I = illuminance

m = magnitude

5.4.2. Stellar Spectral Classes. Under the Harvard system of classification the principal types of spectra are designated by the letters B, A, F, G, K, and M. Stars intermediate to these designations are designated by suffixed numbers from 0 to 9.

The apparent temperatures corresponding to the various spectral classes are not always the same, but vary according to the methods used to measure or calculate the temperature. The following list should be considered only an approximation for main-sequence stars:

Spectral Classification	*Surface Temperature of Star (°K)*
B-0	20,000
A-0	11,000
F-0	7,500
G-0	6,000
K-0	5,000
M-0	3,500

5.4.3. Numbers of Stars. Table 5-2 shows the estimated number of stars brighter than a given magnitude for both photographic and visual magnitudes. From magnitude 0 to 18.5, the figures are based on direct observation; the values from magnitude 18.5 to 21 are extrapolated.

The photographic results are based on all available material such as photographs, star charts, etc. The data for visual magnitudes are derived from the photographic results by allowing for the color of the stars. Very few stars are bluer than class *A*-0, for which class the visual and photographic magnitudes are equal; but many stars are redder and have color indices of +1 magnitude or more. A list of stars brighter visually than the tenth magnitude, for example, will contain many red stars which are photographically of the eleventh magnitude or fainter, and a great many which are photographically fainter than the tenth magnitude. On the other hand, a list of stars

to the tenth photographic magnitude will contain a few blue stars which are visually below the tenth magnitude, but not many. The difference in the numbers in the two columns is thus explained. As seen by the table, this effect increases for the fainter stars, which are generally redder than the brighter ones. Table 5-3 shows the percentage of stars in the six principal spectral classes for various ranges of magnitudes.

TABLE 5-2. ESTIMATED TOTAL NUMBER OF STARS BRIGHTER THAN A GIVEN MAGNITUDE [22]

Photographic Magnitude	*Number of Stars*	
	Photographic	*Visual*
0	–	2
1	–	2
1	–	12
2	–	40
3	–	140
4	360	530
5	1,030	1,620
6	2,940	4,850
7	8,200	14,300
8	22,800	41,000
9	62,000	117,000
10	166,000	324,000
11	431,000	870,000
12	1,100,000	2,270,000
13	2,720,000	5,700,000
14	6,500,000	13,800,000
15	15,000,000	32,000,000
16	33,000,000	71,000,000
17	70,000,000	150,000,000
18	143,000,000	296,000,000
19	275,000,000	560,000,000
20	505,000,000	1,000,000,000
21	890,000,000	—

TABLE 5-3. PERCENTAGE OF STARS OF VARIOUS SPECTRAL CLASSES [22]

Visual Magnitude	*B*-0 *to B*-5	*B*-8 *to A*-3	*A*-5 *to F*-2	*F*-5 *to G*-0	*G*-5 *to K*-2	*K*-5 *to M*-8
< 2.24	28	28	7	10	15	12
2.25 to 3.24	25	19	10	12	22	12
3.25 to 4.24	16	22	7	12	35	8
4.25 to 5.24	9	27	12	12	30	10
5.25 to 6.24	5	38	13	10	28	6
6.26 to 7.25	5	30	11	14	32	7
7.26 to 8.25	2	26	11	16	37	7
8.5 to 9.4	2	18	13	20	36	12
9.5 to 10.4	1	16	12	24	38	9
For All Magnitudes	2	29	9	21	33	6
Photographic Magnitude	*B*-0 *to B*-5	*B*-6 *to A*-4	*A*-5 *to F*-4	*F*-5 *to G*-4	*G*-5 *to K*-4	*K*-5 *to M*-8
8.5 to 9.5	2	31	16	24	24	3
9.5 to 10.5	1	24	16	31	26	3
10.5 to 11.5	1	17	13	40	27	3
11.5 to 12.5	0	10	13	47	26	3
12.5 to 13.5	0	3	10	58	26	2

The data are taken from the publications of the Harvard, McCormick, and Bergedorf Observatories. The discontinuity in trend appearing between the visual and photographic groupings is in accordance with expectations. Of the stars brighter than magnitude 8.5, 99% belong to the six classes listed.

5.4.4. Galactic Concentration of Stars

5.4.4.1. *The Number of Stars (Galactic Concentration) in Different Parts of the Sky.* Table 5-4 shows the number of stars per square degree brighter than a given photographic magnitude, for different galactic latitudes.

TABLE 5-4. NUMBER OF STARS PER SQUARE DEGREE BRIGHTER THAN PHOTOGRAPHIC MAGNITUDE AS A FUNCTION OF GALACTIC LATITUDES [22]

Photographic Magnitude	+90°	+40°	+20°	+10°	0°	−10°	−20°	−40°	−90°
5.0	0.014	0.0175	0.023	0.031	0.059	0.045	0.032	0.0178	0.012
6.0	0.039	0.053	0.071	0.089	0.166	0.126	0.087	0.051	0.042
7.0	0.015	0.151	0.20	0.257	0.436	0.323	0.224	0.144	0.123
8.0	0.275	0.42	0.59	0.741	1.230	0.851	0.617	0.398	0.316
9.0	0.724	1.12	1.62	2.14	3.55	2.34	1.69	1.10	0.832
10.0	1.78	2.95	4.50	5.89	10.5	6.61	4.68	2.95	2.09
11.0	4.3	7.4	12.0	16.2	30.9	18.2	12.8	7.76	5.25
12.0	10.2	18.2	32.0	43.6	89.1	50.1	34.7	19.50	13.2
13.0	24.0	43.0	79.0	112.0	245.0	138.0	89.1	47.8	30.2
14.0	50.0	93.0	190.0	282.0	661.0	371.0	218.0	107.0	60.3
15.0	95.0	200.0	457.0	708.0	1660.0	977.0	525.0	218.0	104.0
16.0	182.0	407.0	1047.0	1778.0	3981.0	2455.0	1175.0	436.0	182.0
17.0	338.0	794.0	2291.0	4365.0	9120.0	5754.0	2512.0	832.0	302.0
18.0	616.0	1413.0	4677.0	9330.0	20890.0	12590.0	4786.0	1514.0	501.0
19.0	770.0	2180.0	6860.0	–	–	–	–	–	–
20.0	–	–	–	–	–	–	–	–	–
21.0	1670.0	5000.0	21200.0	–	–	–	–	–	–

5.4.4.2. *Galactic Concentration of Stars of Various Spectral Classes.* Table 5-5 shows the average number of stars per 100 square degrees near the galactic equator and in regions remote from it for the six principal spectral classes.

An approximation to the number of stars of a certain spectral class and magnitude range can be obtained by applying the data of Table 5-3 to Table 5-2, since Table 5-2 gives the estimated number of stars brighter than a given magnitude for each magnitude. For example, by interpolation of Table 5-2, the estimated number of stars brighter than magnitudes 7.25 and 8.25 may be obtained. By subtraction, the number of stars in the magnitude range 7.25 to 8.25 is obtained. The percentage of stars of the six principal spectral classes for this range of magnitudes, as shown in Table 5-3, can be used to obtain the approximate number of stars in these spectral classes for this range of magnitudes.

TABLE 5-5. GALACTIC CONCENTRATION OF STARS OF THE PRINCIPAL SPECTRAL CLASSES IN 100 SQUARE DEGREES NEAR GALACTIC EQUATOR [22]

Stellar Magnitudes	*Galactic Latitudes*	*B*	*A*	*F*	*G*	*K*	*M*	*Total*
Above $7^m.0$								
	40°–90°	0.2	6.6	3.0	3.4	10.2	1.5	24.9
	0°	10.8	21.1	5.1	5.1	15.1	3.9	61.1
$7^m.0$ to $8^m.25$								
	40°–90°	0.1	6.6	9.5	16.4	32.8	6.1	71.5
	0°	18.9	75.8	13.6	20.9	53.9	13.6	196.7

Table 5-6 gives more detailed information of the distribution of stars by spectral class and magnitude. There are differences in the data of Table 5-5 and 5-6 because somewhat different areas of the sky were considered in preparing the tables. For example, Table 5-5 considers the latitude from 40° to 90°, whereas Table 5-6 considers

TABLE 5-6. GALACTIC CONCENTRATION OF STARS OF VARIOUS SPECTRAL CLASSES [22]

Spectrum Visual Magnitude	*B*	*A*	*F*	*G*	*K*	*M*
			Galactic Latitude 0° *to* 5°			
< 6.0	4.5	6.0	1.7	2.1	3.5	1.3
6.0 to 7.0	6.3	15	3.4	3.0	12	2.6
7.0 to 8.25	19	76	14	21	54	14
8.5 to 9.4	46	190	85	96	200	57
9.5 to 10.4	82	610	240	310	490	150
Photographic Magnitude						
9.5 to 10.5	38	510	150	220	180	19
10.5 to 11.5	87	970	430	720	460	42
11.5 to 12.5	100	1390	1200	1960	940	140
Visual Magnitude			*Galactic Latitude* 60° *to* 90°			
< 6.0	0.2	2.6	0.8	1.0	2.9	0.7
6.0 to 7.0	0	3.8	1.8	2.4	7.5	0.7
7.0 to 8.25	0	7.4	9.2	16	32	6.3
8.5 to 9.4	0	8	20	83	75	0
9.5 to 10.4	0	8	20	170	210	16
Photographic Magnitude						
9.5 to 10.5	0	9	32	120	75	9
10.5 to 11.5	0	10	27	290	160	12
11.5 to 12.5	0.9	14	34	680	270	26

Note: The data are taken from the publications of the Harvard, McCormick, and Bergedorf Observatories.

TABLE 5-7. INDEX OF APPARENT GALACTIC CONCENTRATION [22]

Visual Magnitude	*B*	*A*	*F*	*G*	*K*	*M*
< 6.0	22	2.8	2.3	2.1	1.2	1.9
6.0 to 7.0	–	4.0	1.9	1.2	1.5	3.7
7.0 to 8.25	–	10	1.5	1.3	1.7	2.2
8.5 to 9.4	–	24	4.2	1.2	2.7	–
9.5 to 10.4	–	76	12	1.8	2.3	0.9
Photographic Magnitude						
9.5 to 10.5	–	56	4.8	1.8	2.4	2.1
10.5 to 11.5	–	97	16	2.5	2.9	3.5
11.5 to 12.5	–	99	35	2.9	3.5	5.5

Note: The irregularities here are attributable in part to inadequate sampling.

the latitude from 60° to 90°, in arriving at an average galactic distribution. The most important difference is that Table 5-6 has been prepared by selecting narrower ranges of stellar magnitude.

Data of galactic distribution are not presented for stars of magnitudes less than 5 because the total number of these stars is not large enough to make the concept of the number of stars per square degree meaningful.

Table 5-7, an index of apparent galactic concentration, has been prepared from Table 5-6 by taking the ratios of numbers of stars in low latitudes to the numbers in high latitudes. For a given spectral class, more stars are concentrated in the lower galactic latitudes as the index number becomes higher.

5.4.5. Spectral Distribution of Stellar Radiation.* Figure 5-21 shows relative spectral distribution of stellar radiation as a function of star classes and surface temperature. The family of curves in Fig. 5-22 shows absolute spectral distribution of

*See also *Notes Added in Proof*, page 171.

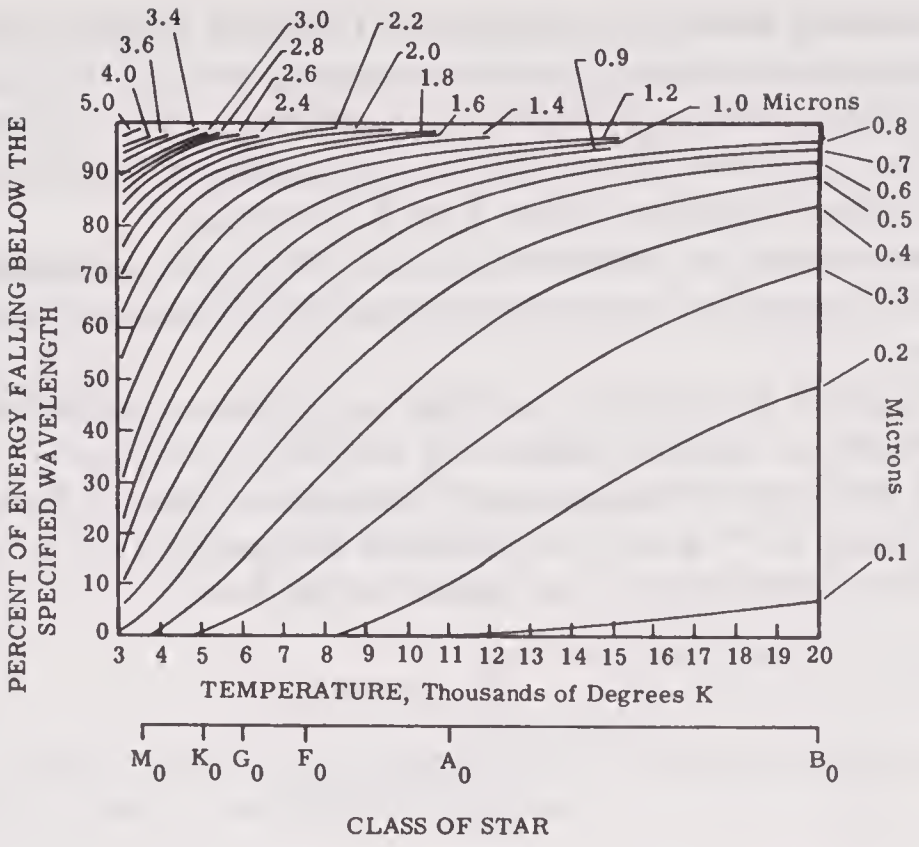


FIG. 5-21. Relative spectral distribution of stellar radiation as a function of star classes [22].

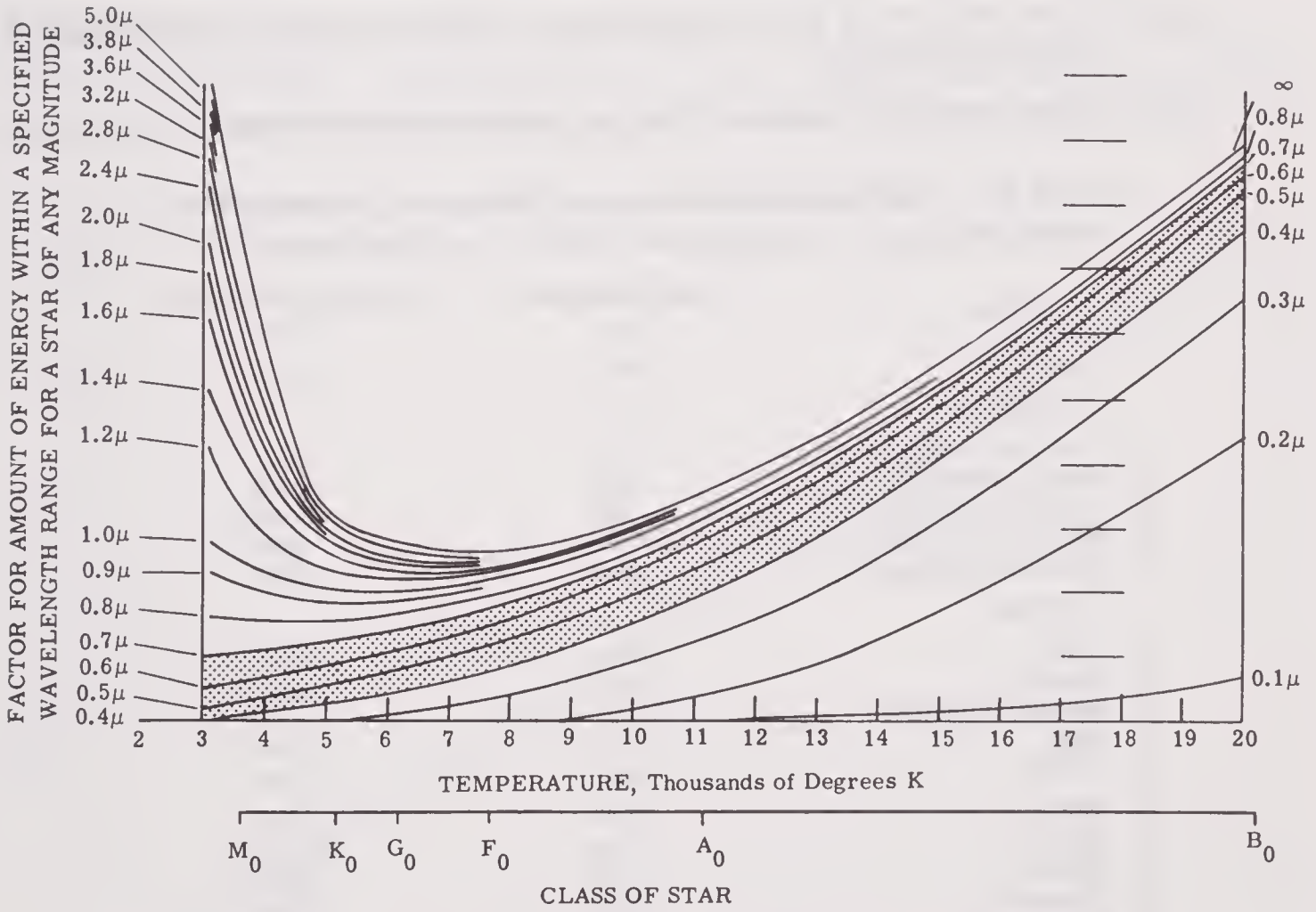


FIG. 5-22. Absolute spectral distribution of stellar radiation [22].

stellar radiation. In this figure, the absolute magnitude of the radiant energy falling below a specified wavelength is plotted as a function of the surface temperature of the stars. Further, the curves have been normalized so that the amount of energy in the visible region is constant for all the stars of any given magnitude. This value is represented by one vertical division of the graphic scale on Fig. 5-22.

5.4.6. Determining Spectral Irradiance of Celestial Bodies [23]. The following data and methods permit determining with reasonable accuracy the spectral irradiance values of the brightest stars and planets in the infrared region. These data have been calculated from published measurements of visible irradiance and effective temperature, and include the complete spectral region from 0.1 to 100 μ.

The data used pertain to irradiance received above the atmosphere. Values for absorption by the atmosphere in the various spectral regions can be readily applied to the chart values.

Table 5-8 [21] shows the visible magnitude and effective temperature ($T_{eff} = W/\sigma$) values for the brightest celestial bodies and also for the important "red stars." The list contains all the stars in Schlessinger's *Catalogue of Bright Stars* which give an irradiance of at least 10^{-12} w cm^{-2} in either the PbS region (1–3 μ) or the bolometer region (0.3–13.5 μ). Equation (5-2) is plotted in Fig. 5-23.

$$\eta_e\,(T) = \frac{\int_0^\infty J_\lambda(T) S_{e\lambda}\, d\lambda}{\int_0^\infty J_\lambda(T)\, d\lambda} \tag{5-2}$$

where $\eta_e(T)$ is the fraction of total radiation emitted by a blackbody at some temperature T, visible to the standard observer

$J_\lambda(T)$ is the ordinate of the Planck blackbody radiation curve at wavelength λ and temperature T

$S_{e\lambda}$ is the fractional response of the eye at the same wavelength.

TABLE 5-8. VISUAL MAGNITUDES AND EFFECTIVE TEMPERATURE OF PLANETS AND THE BRIGHTEST VISUAL AND RED STARS [21]

	Name	*Visual Magnitude* (m_v)	*Effective Temperature* T (°K)
1.	Moon (full)	−12.2	5,900
	(Planets)		
2.	Venus (at brightest)	−4.28	5,900
3.	Mars (at brightest)	−2.25	5,900
4.	Jupiter (at brightest)	−2.25	5,900
5.	Mercury (at brightest)	−1.8	5,900
6.	Saturn (at brightest)	−0.93	5,900
	(Stars)		
1.	Sirius	−1.60	11,200
2.	Canopus	−0.82	6,200
3.	Rigel Kent (double)	0.01	4,700
4.	Vega	0.14	11,200
5.	Capella	0.21	4,700
6.	Arcturus	0.24	3,750
7.	Rigel	0.34	13,000
8.	Procyon	0.48	5,450
9.	Achernar	0.60	15,000
10.	β Centauri	0.86	23,000
11.	Altair	0.89	7,500
12.	Betelguex (variable)	0.92	2,810
13.	Aldebaran	1.06	3,130
14.	Pollux	1.21	3,750
15.	Antares	1.22	2,900
16.	α Crucis	1.61	2,810
17.	Mira (variable)	1.70	2,390
18.	β Gruis	2.24	2,810
19.	R. Hydrae (variable)	3.60	2,250

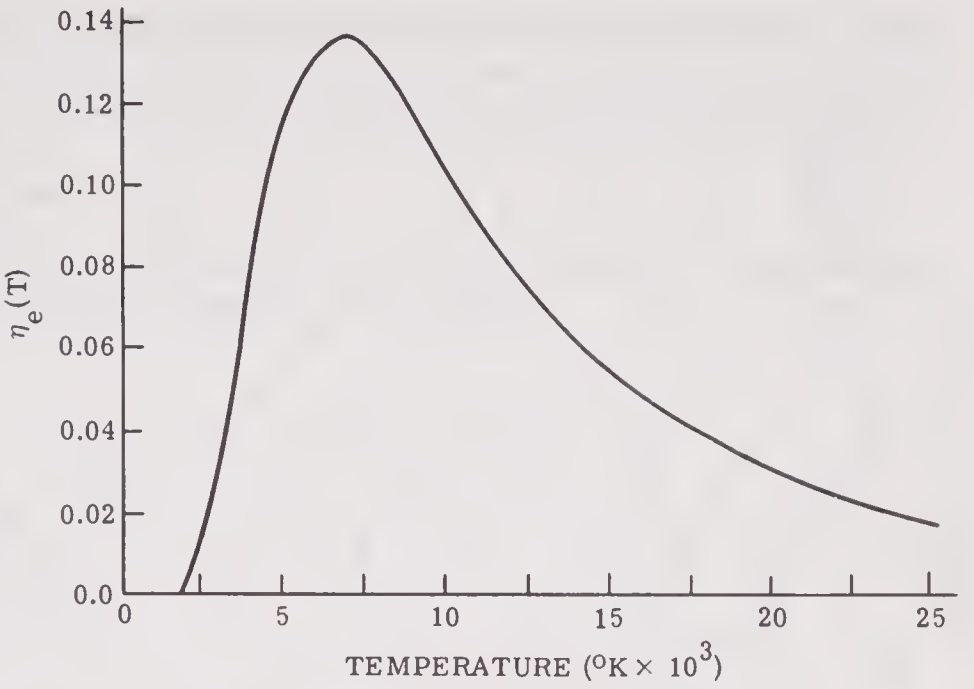


FIG. 5-23. Fraction of the total radiation emitted by a blackbody at temperature T, visible to the standard observer [21].

After $\eta_e(T)$ is found, the stellar magnitude m_v of the body may be used to obtain the total blackbody spectral irradiance, as follows:

$$m_v = 2.5 \log_{10} \frac{I(m_v)}{I_0} \tag{5-3}$$

At the top of the atmosphere, zero visible magnitude corresponds to a visible irradiance, I_0, of 3.1×10^{-13} w/cm². The value for $I(m_v)$, for any quoted value of stellar magnitude, may then be obtained by the solution of Eq. (5-3). This function is plotted in Fig. 5-24.

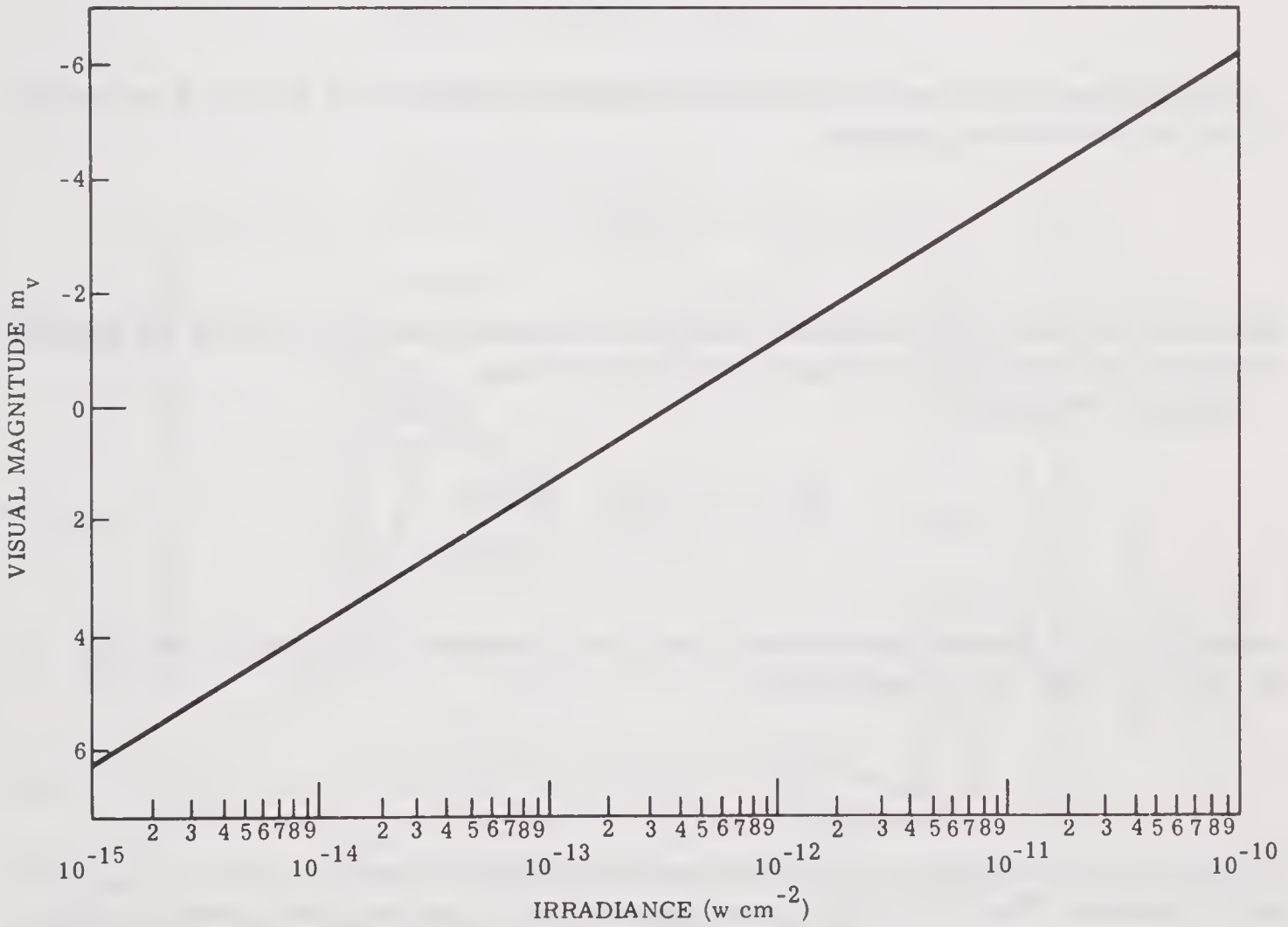


FIG. 5-24. Effective irradiance in the visible-region (standard observer) *versus* visual magnitude [21].

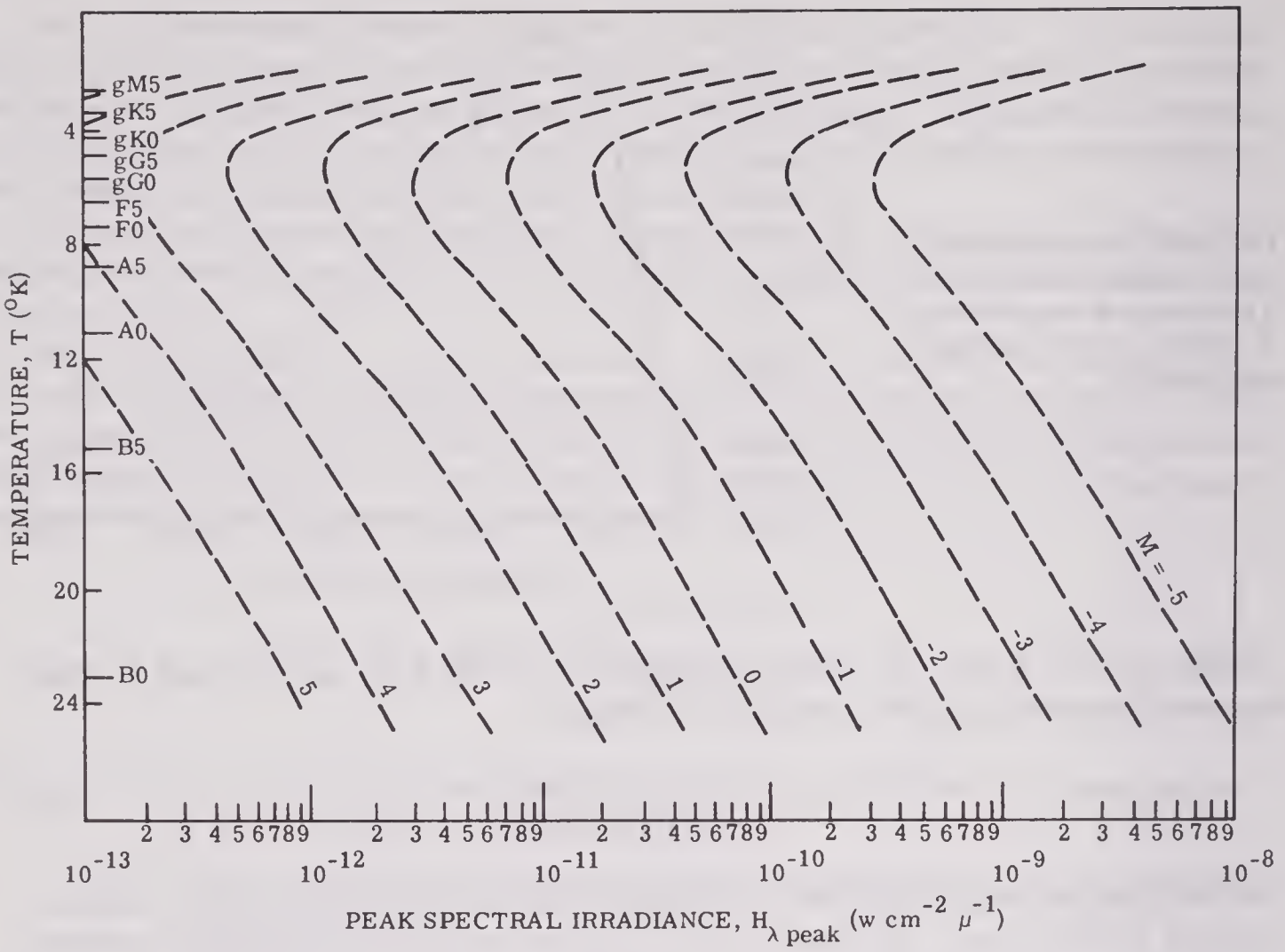


FIG. 5-25. Peak spectral irradiance from values of visual magnitude and effective temperature or spectral class [21].

The irradiance received over the total wavelength spectrum at the top of the atmosphere is therefore the quantity

$$\frac{I(m_v)}{\eta_e(T)}$$

and once the value of peak spectral irradiance is determined, the shape of the spectral irradiance curve follows the Planck radiation function.

The peak irradiance is

$$H_{\lambda\ peak} = \frac{I(m_v)}{\eta_e(T)} \times \frac{W_{\lambda\ max}}{\int_0^\infty W\, d_\lambda} \tag{5-4}$$

where $W_{\lambda\ max}$ is the maximum value of the Planck function, and equals $1.290 \times 10^{-15} T^5$ w cm^{-2} μ^{-1}. Eq. (5-4) then becomes

$$H_{\lambda\ peak} = \frac{I(m_v)}{\eta_e(T)} \times 2.272T \times 10^{-4}\ \text{w cm}^{-2}\ \mu^{-1} \tag{5-5}$$

Equation (5-5) evaluated and plotted as a function of T for various values of magnitude m_v is shown in Fig. 5-25. This graph can be used to find the peak spectral radiance, $H_{\lambda\ peak}$, for any values of T and m_v. Using Fig. 5-25 and Wien's law, the spectral irradiance curves for any star or planet may be obtained. (In determining the spectral

irradiance of the planets, an effective temperature of 5900°K was assumed.) The shape of all these irradiance curves are identical; they are blackbody curves normalized to their peak value.

5.5 The Earth as a Background*

5.5.1. Geometric Relationships. Figures 5-26, 5-27, and 5-28 present some important relationships bearing on satellite viewing of the earth.

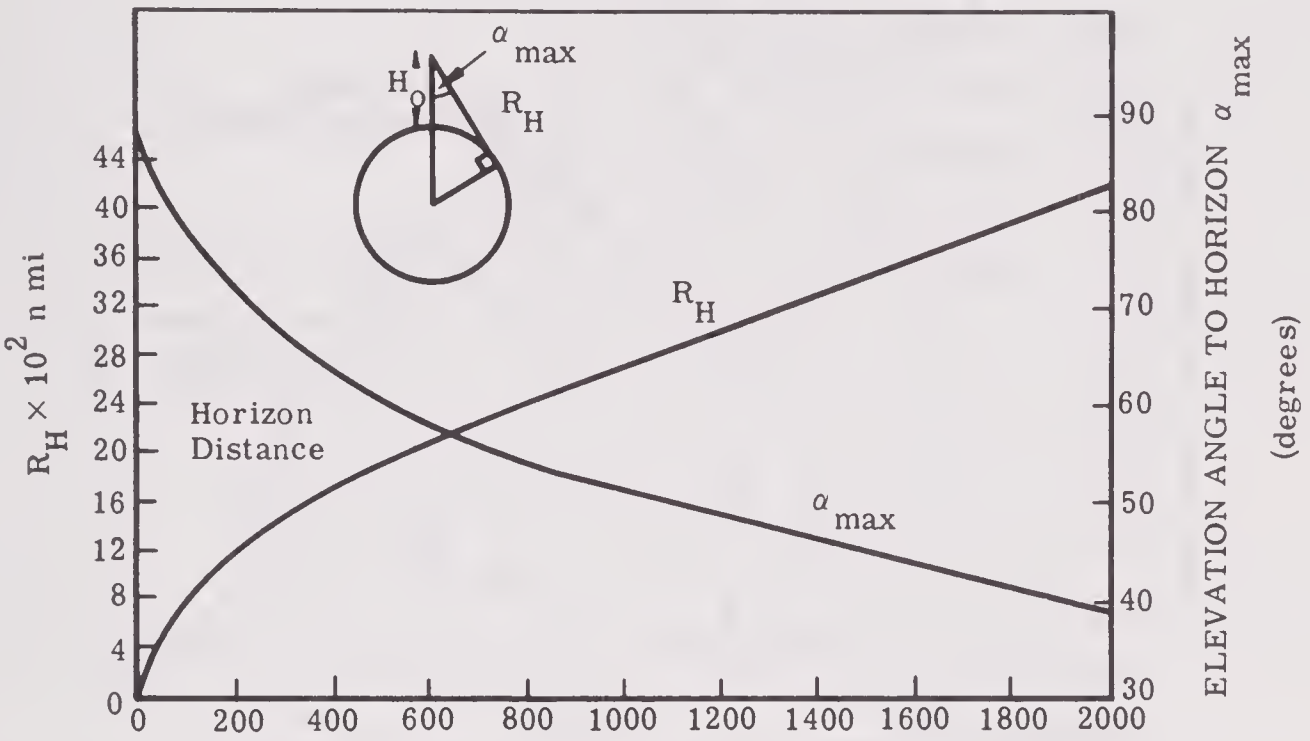


FIG. 5-26. Ranges and view angles [3].

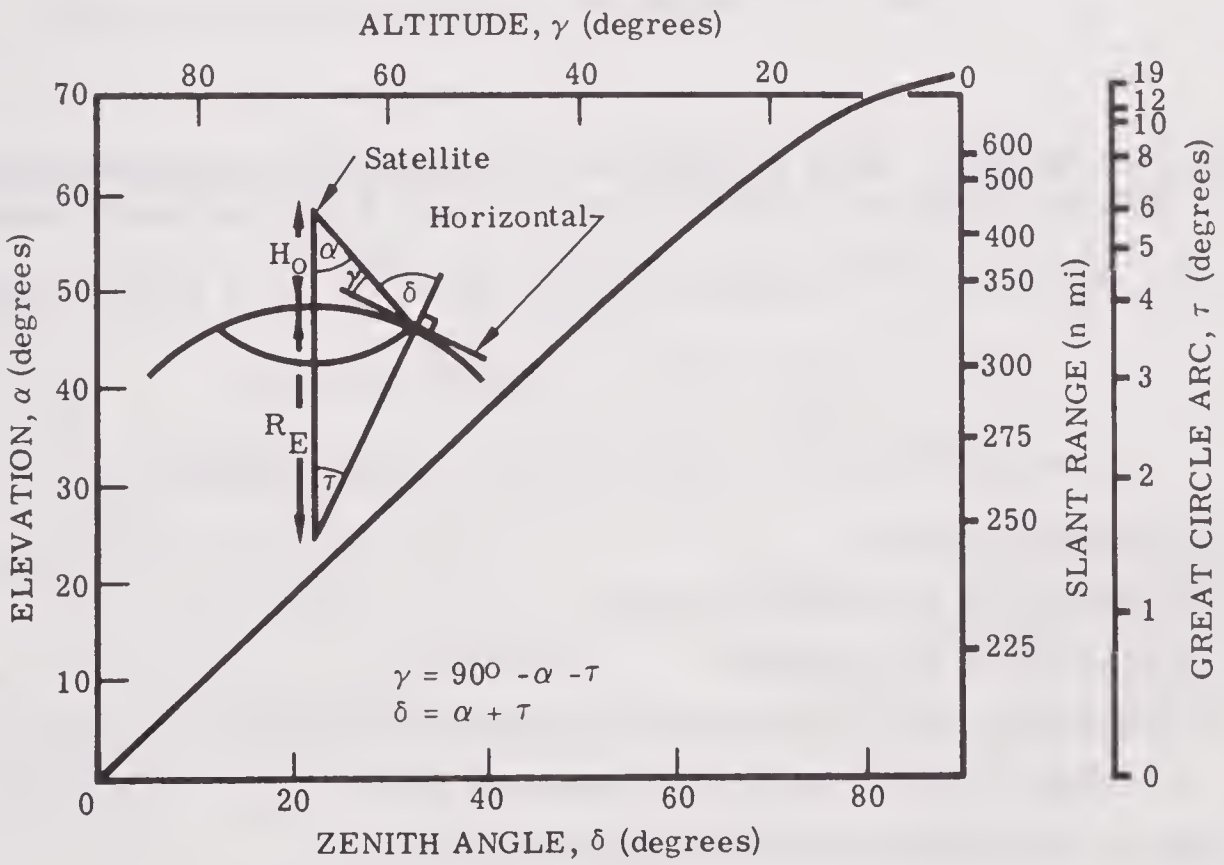


FIG. 5-27. View angles for 200 n mi orbit [3].

*See also *Notes Added in Proof*, page 171.

In Fig. 5-28, point P (as an example) represents a vehicle 400 n mi high at an elevation of 35°. The slant range is 655 n mi, and the great circle arc (angle between vectors located at the center of the earth and pointing respectively to the satellite and to the ground point viewed) is 7°.

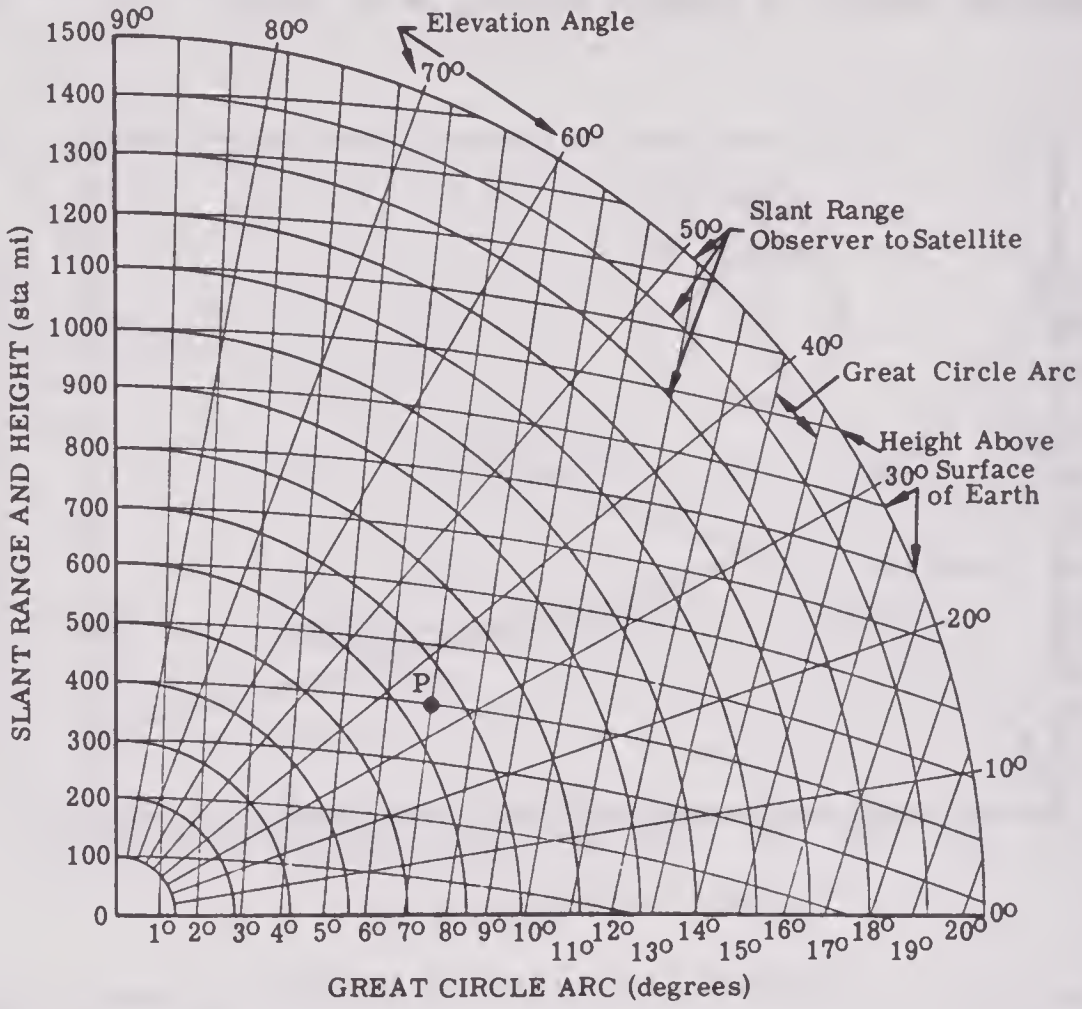


FIG. 5-28. Satellite coordinate conversion [3].

5.5.2. Path Lengths. Refer to Fig. 5-29. The length of a line between any two points at different altitudes is found according to the following general equations:

$$(R_E + C)^2 = (R_E + A)^2 + L^2 + 2L(R_E + A)\cos\delta = (R_E + A)^2 + L^2 + 2L(R_E + A)\sin\gamma$$

$$L^2 = (R_E + C)^2 - (R_E + A)^2 - 2L(R_E + A)\sin\gamma$$

$$L = \sqrt{(R_E + C)^2 - (R_E + A)^2\cos^2\gamma} - (R_E + A)\sin\gamma \qquad (5\text{-}6)$$

where R_E = radius of sphere

A = altitude of the background point

C = altitude of the observer

γ = elevation angle of the background point position (from local horizontal)

α = elevation angle from nadir at observer's point.

The angle α is computed by the following relationship:

$$\alpha = \cos^{-1}\frac{(R_E + C)^2 + L^2 - (R_E + A)^2}{2(R_E + C)L} \qquad (5\text{-}7)$$

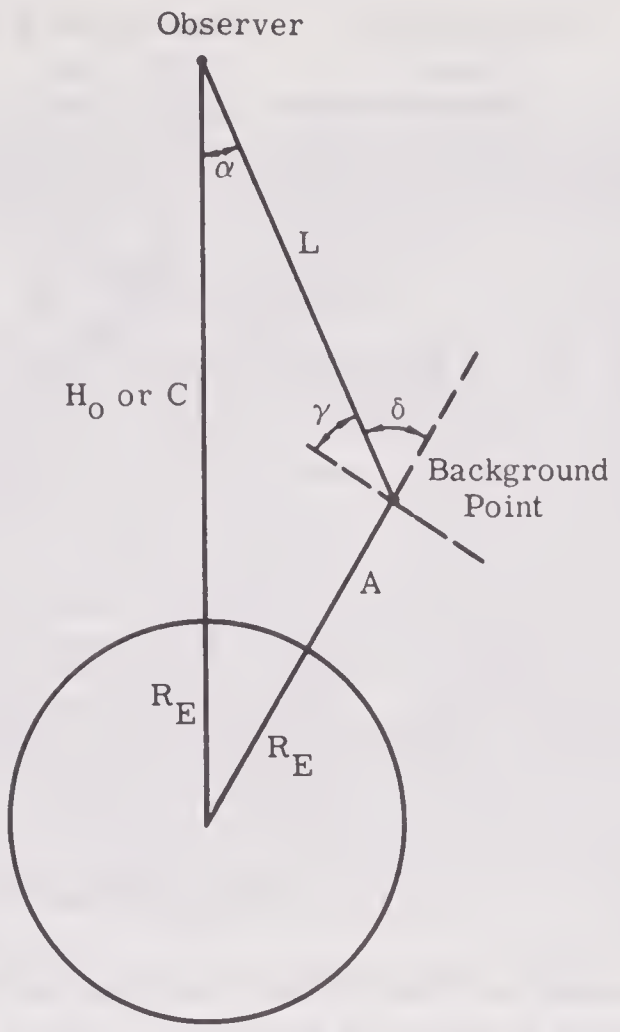


FIG. 5-29. Path length diagram.

Figure 5-30 illustrates the scattering angle β, the sun's elevation angle γ, the satellite scanner's elevation angle α, and the scanner's azimuth angle ϕ from the direction of the sun.

The scattering angle is:

$$\beta = \cos^{-1} (\cos \gamma \sin \alpha \cos \phi - \sin \gamma \cos \alpha) \tag{5-8}$$

Figure 5-31 illustrates these angles for a spherical earth. Equation (5-8) becomes

$$\beta = \cos^{-1} [\cos (90° - (\xi - \Delta)] \sin \alpha \cos \phi - \sin [90° - (\xi - \Delta)] \tag{5-9}$$

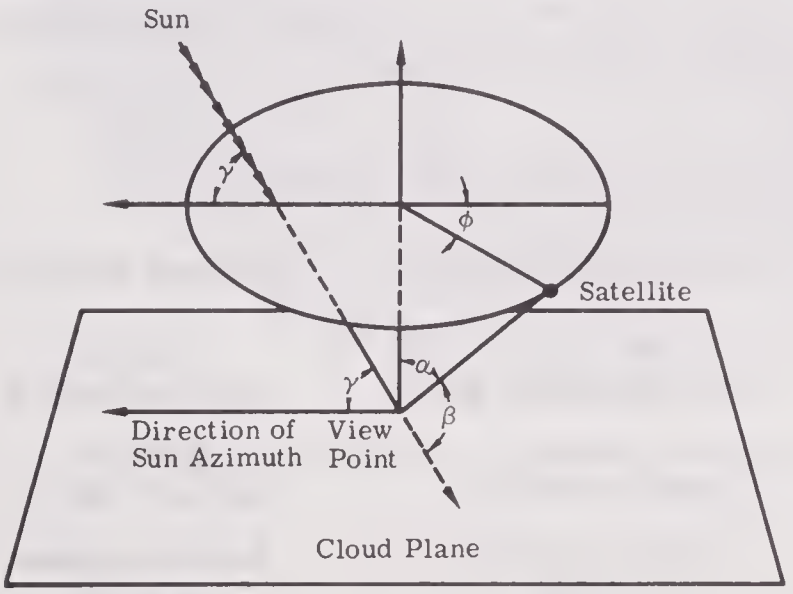


FIG. 5-30. Scattering-angle geometry [3].

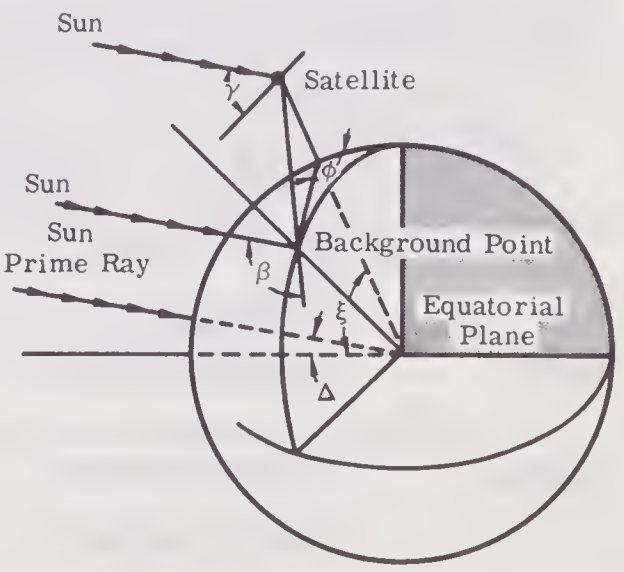


FIG. 5-31. Solar scattering angle [3].

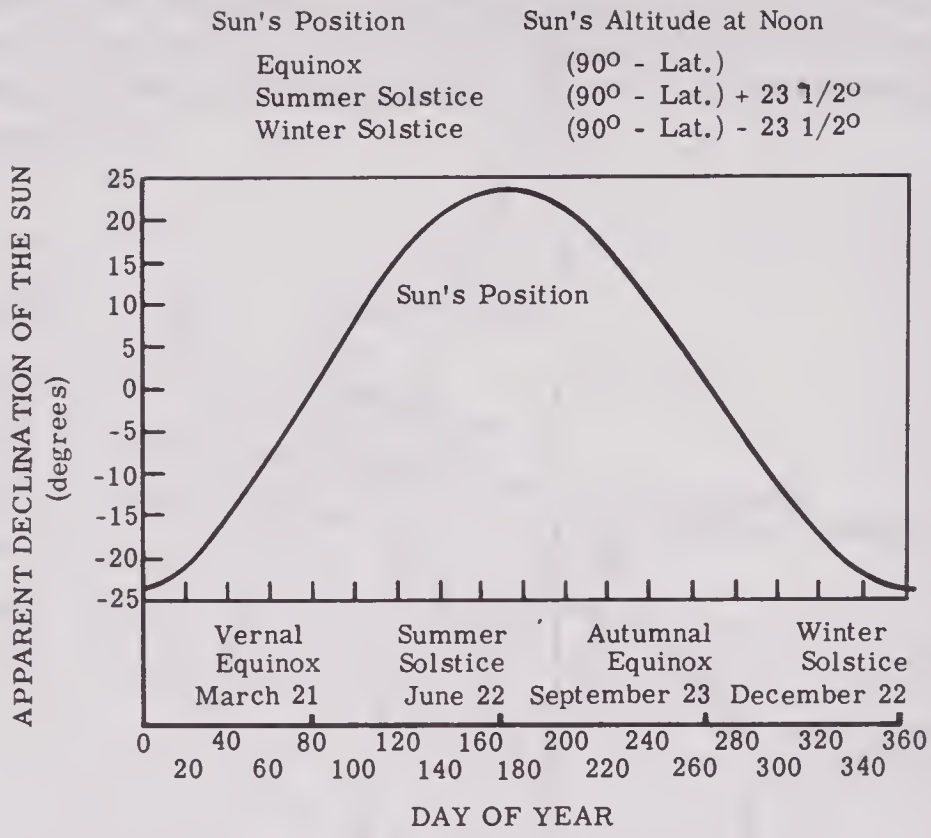


FIG. 5-32. Solar declination to equator [3].

where ξ is the elevation angle (latitude) of the observed point from the earth's equatorial plane, and Δ is the latitude of the sun's prime ray. Figure 5-32 shows the range of the sun's declination angle. (Both ξ and Δ are taken positive for a North latitude.)

5.6. Cloud Meteorology

Clouds are classified into ten main groups called genera. These are cirrus, cirrocumulus, cirrostratus, altocumulus, altostratus, nimbostratus, stratocumulus, stratus, cumulus, and cumulonimbus.

The part of the atmosphere in which clouds are usually present is divided into three regions. Each region is defined by the range of levels at which clouds of certain genera occur most frequently.

(*a*) High-level clouds—cirrus, cirrocumulus, and cirrostratus

(*b*) Middle-level clouds—altocumulus

(*c*) Low-level clouds—stratocumulus and stratus

The regions overlap, and their limits vary with latitude. Their approximate ranges are shown in Table 5-9. Figures 5-33 to 5-38 show the mean cloudiness in percentage of sky cover throughout the world for various months of the year.

TABLE 5-9. DEFINITION OF CLOUD STATE ALTITUDES [3]

Cloud Level	*Polar Regions*	*Temperate Regions*	*Tropical Regions*
High	3-8 km (10,000-25,000 ft)	5-13 km (16,500-45,000 ft)	6-18 km (20,000-60,000 ft)
Middle	2-4 km (6500-13,000 ft)	2-7 km (6500-23,000 ft)	2-8 km (6500-25,000 ft)
Low	Earth's surface to 2 km (6500 ft)	Earth's surface to 2 km (6500 ft)	Earth's surface to 2 km (6500 ft)

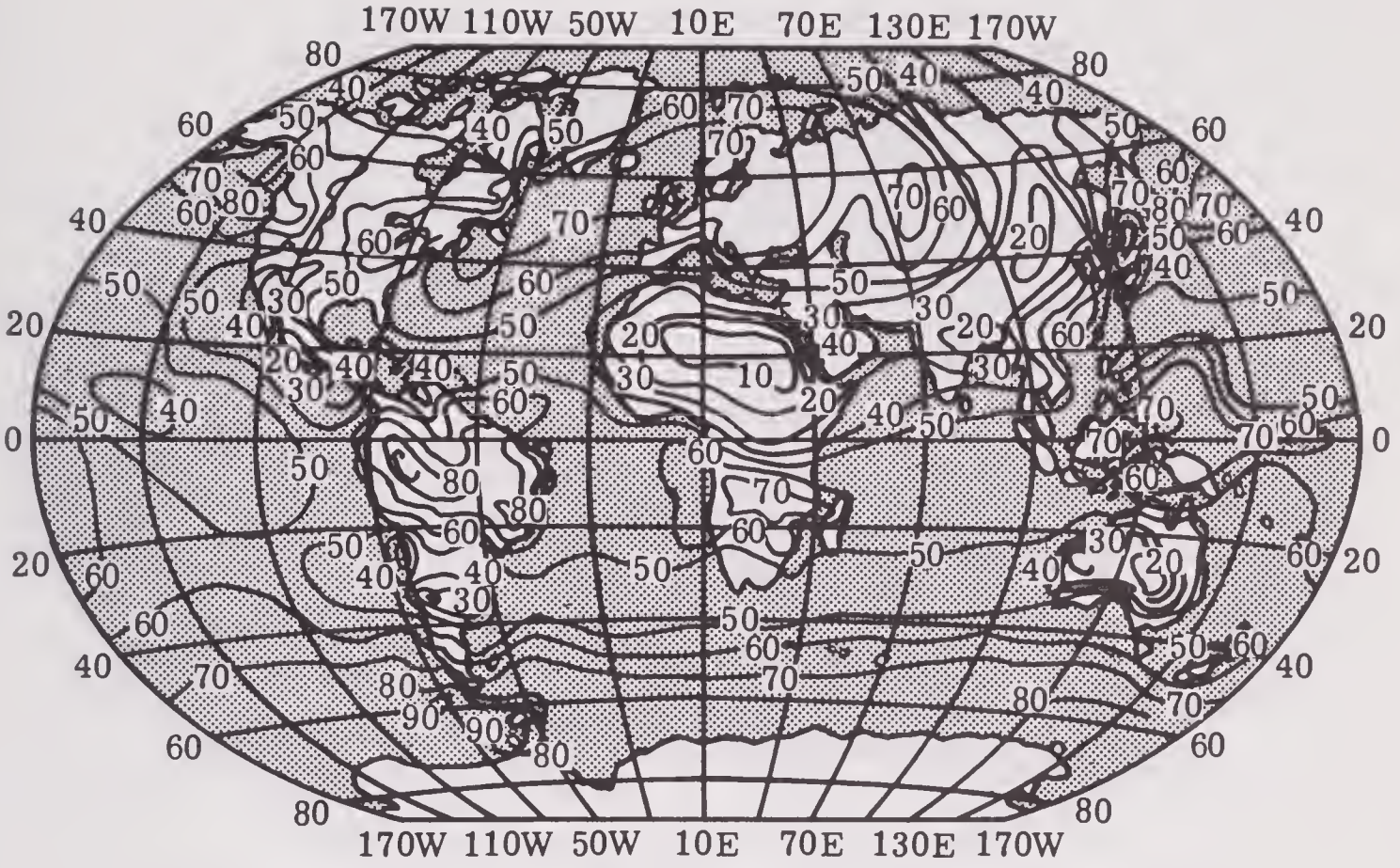


FIG. 5-33. Mean cloudiness in percentage of sky cover, month of January [25].

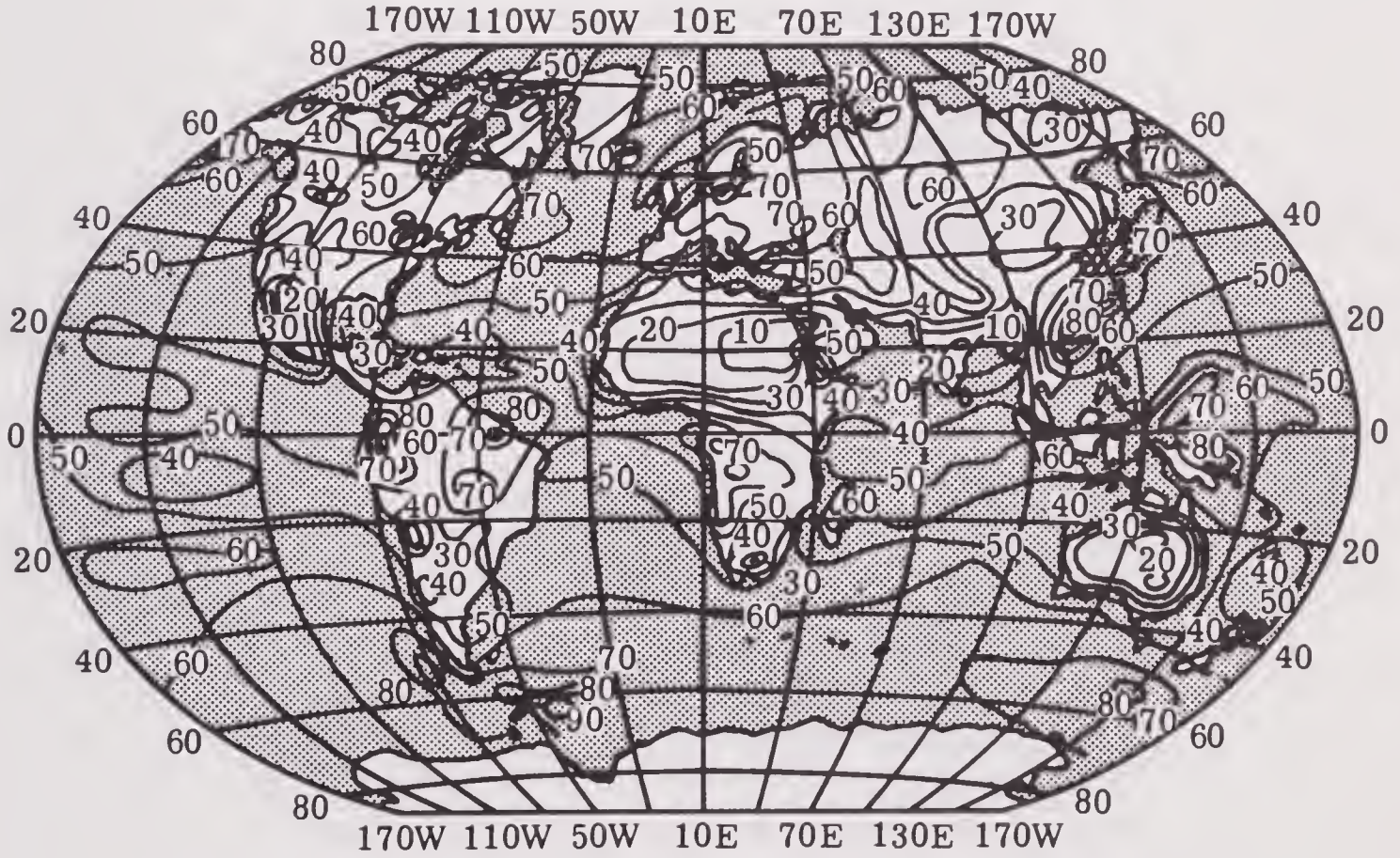


FIG. 5-34. Mean cloudiness in percentage of sky cover, month of March [25].

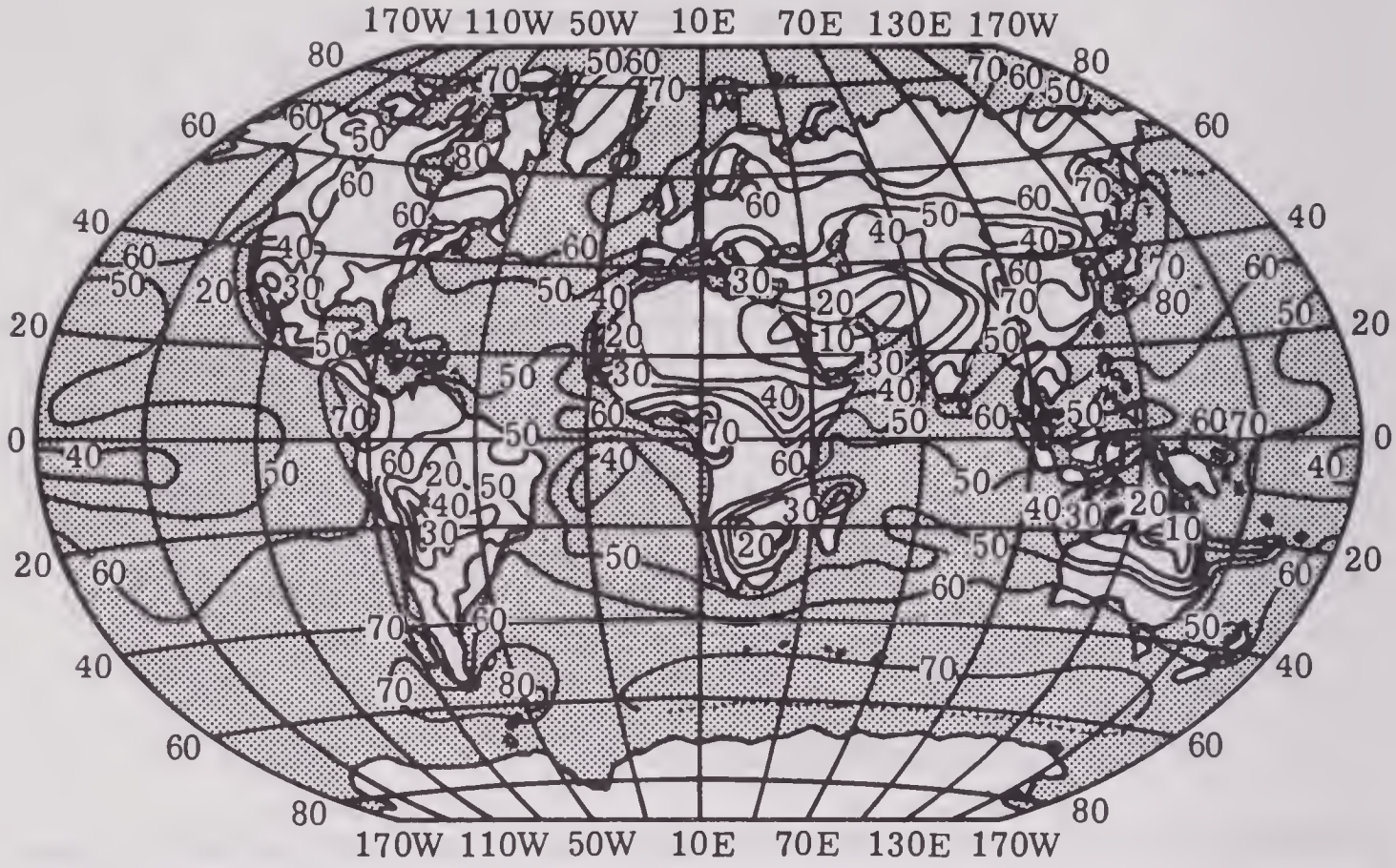


FIG. 5-35. Mean cloudiness in percentage of sky cover, month of May [25].

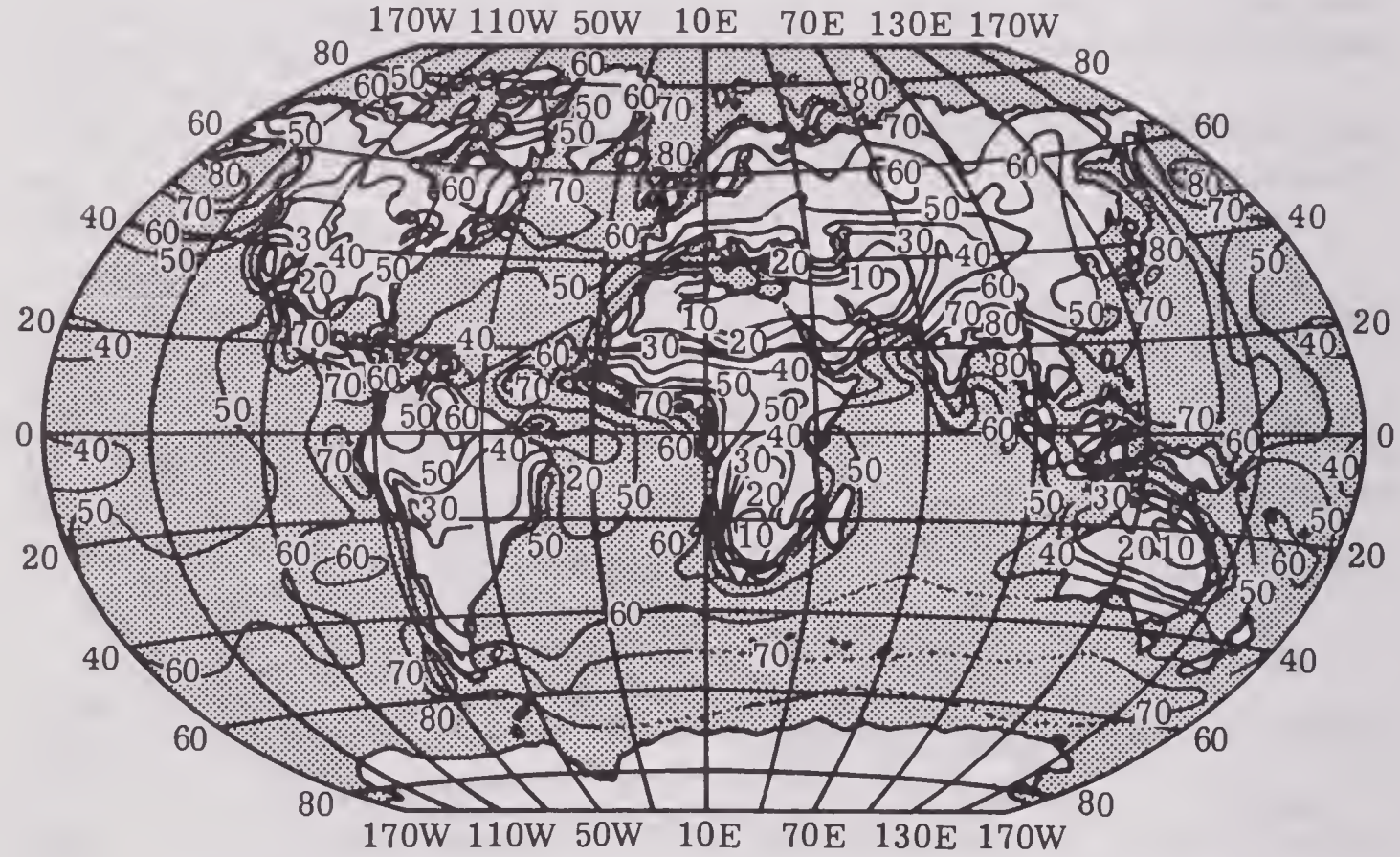


FIG. 5-36. Mean cloudiness in percentage of sky cover, month of July [25].

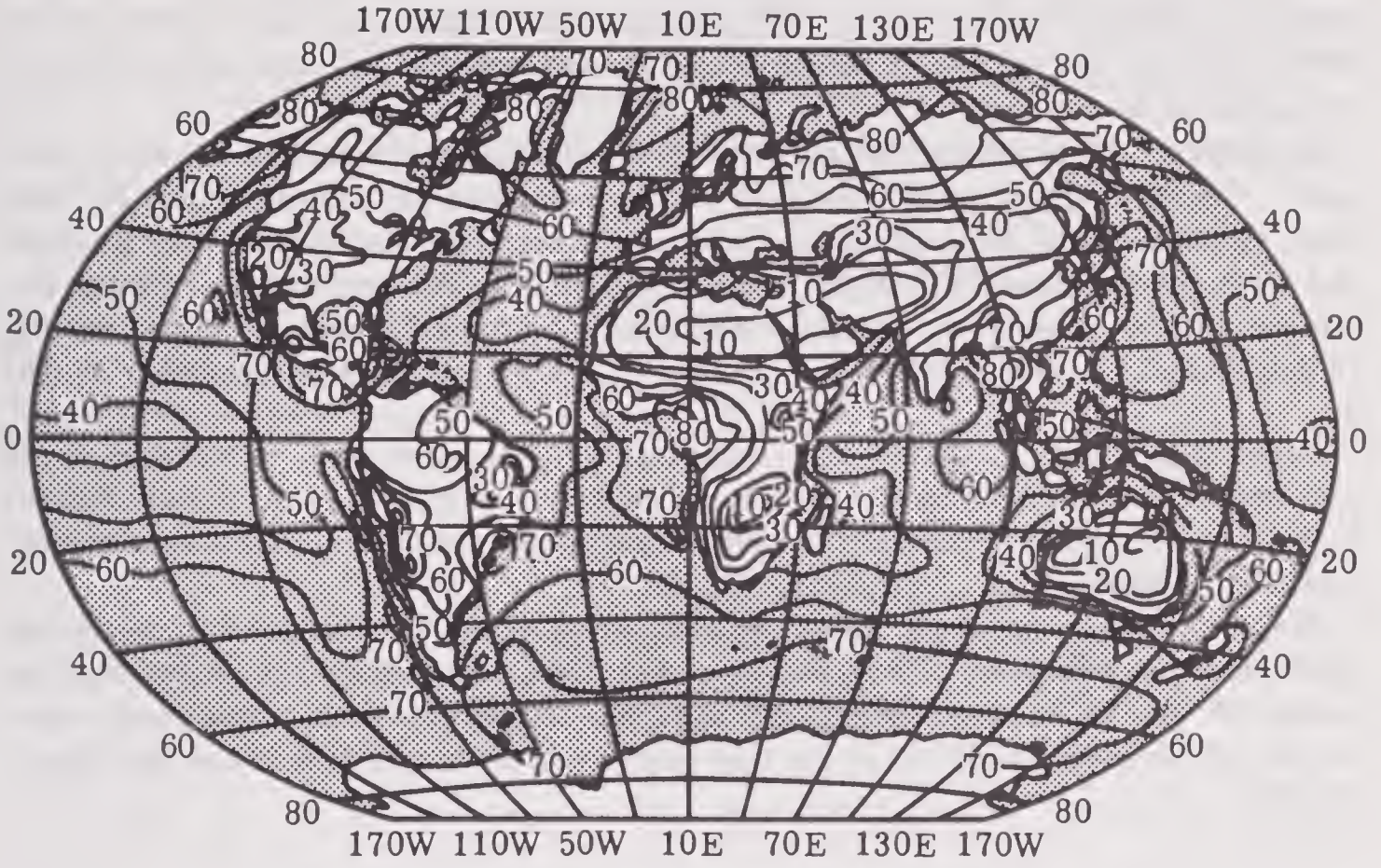


FIG. 5-37. Mean cloudiness in percentage of sky cover, month of September [25].

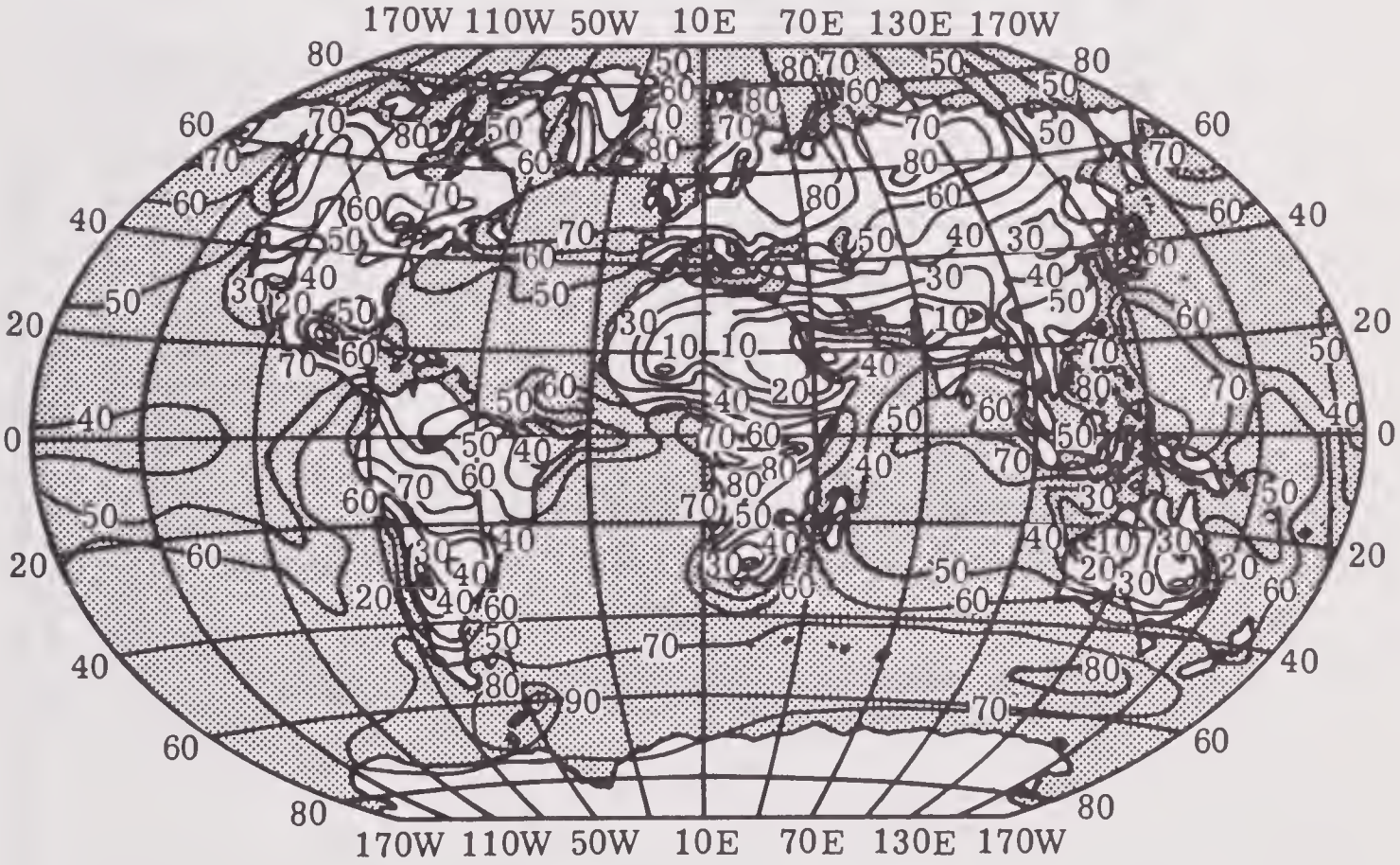


FIG. 5-38. Mean cloudiness in percentage of sky cover, month of November [25].

5.6.1. Cirrus Clouds. The tropopause represents the upper limit of the cloud atmosphere. The highest clouds appearing within the troposphere are composed of large ice crystals of about 100 μ. Frequently these particles become oriented in the same direction, giving rise to unusual visible, and possibly infrared, effects such as haloes and arcs.

Tropopause and cloud top statistics are not available for the central Eurasian land mass. Cirrus height observations have not been reported anywhere north of 55° latitude. Inferences can be made about the annual tropopause distribution over Eurasia, and from this a cirrus top height model constructed. The correlation between the two parameters is based on American statistics. Between 50° and 70°N it is expected that 90% of the annual clouds will be below 32,000 ft, and 99% will be below 36,000 ft.

Figure 5-39 shows cloud top and tropopause heights based on a collation of cirrus and tropopause data averaged on a yearly basis for the entire United States. Figures 5-40 and 5-41 represent the distribution of tropopause and cloud heights between 50° and 90°N latitude.

Based on deductions from Asian climatology, a crude time-frequency occurrence chart has been estimated and is shown in Fig. 5-42. Averaging the entire year between 50° and 70°N, cirrus clouds are expected 35% of the time. This means that cirrus will be encountered 1% of the time above 34,000 ft, and 10% of the time above 30,000 ft.

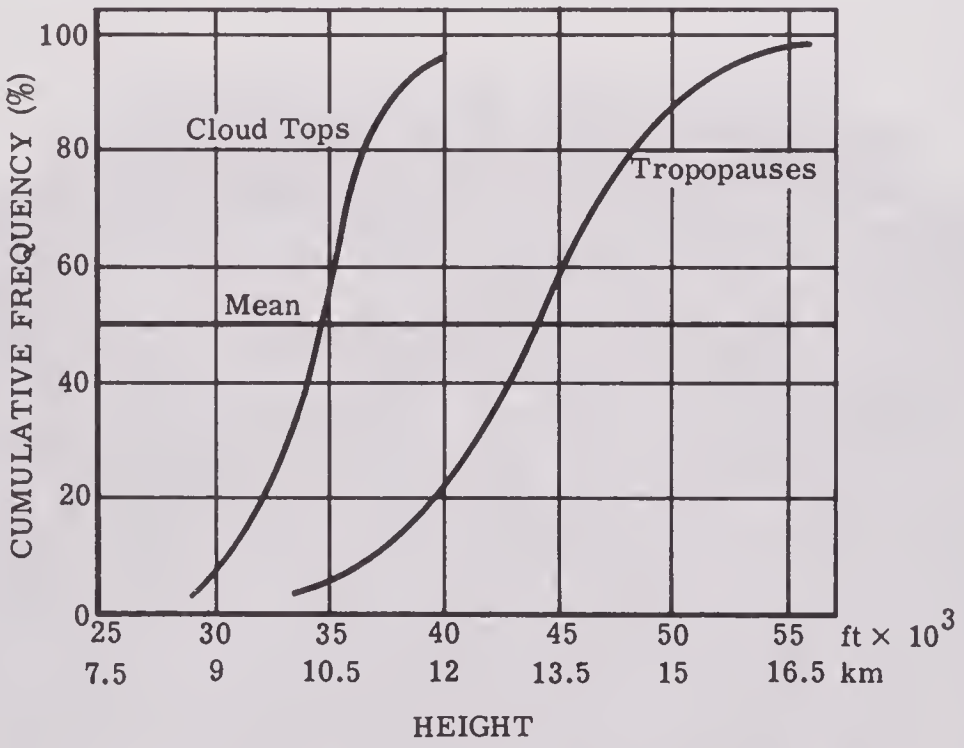


FIG. 5-39. Distribution of cloud top and tropopause heights, United States average [3].

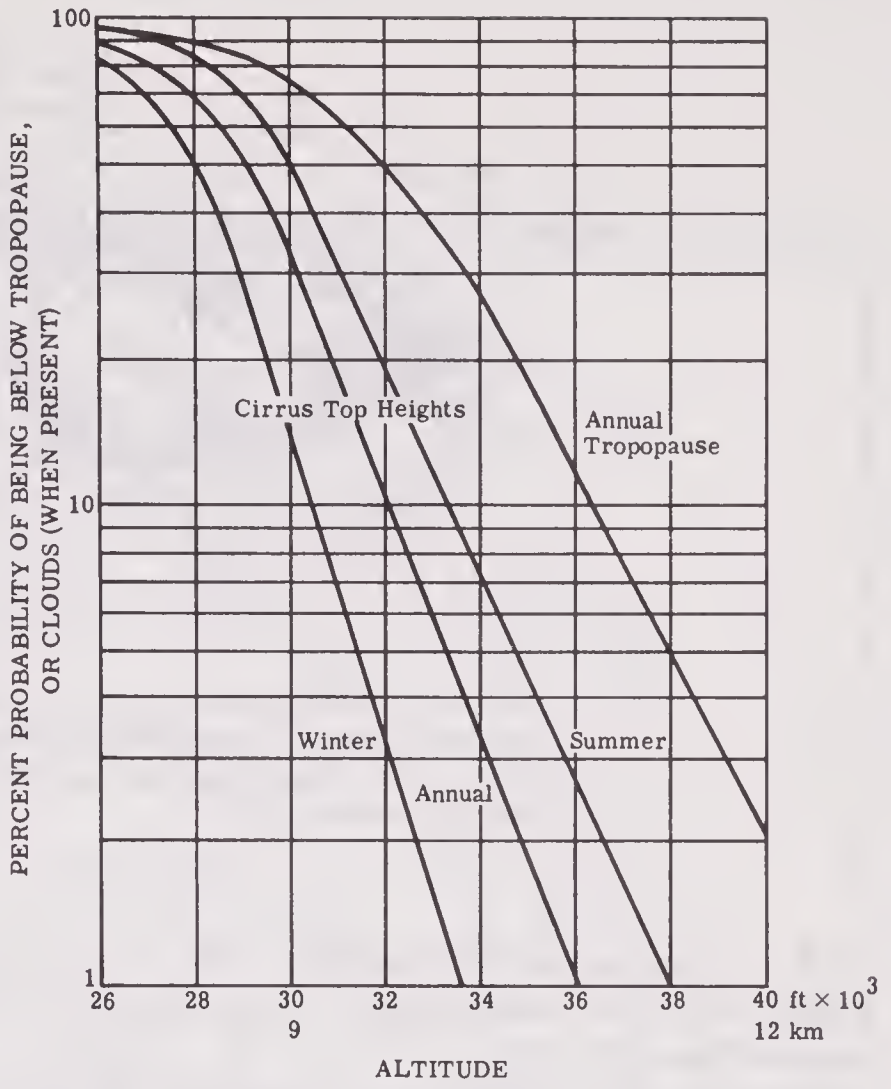


FIG. 5-40. Distribution of tropopause and cloud tops, 50° to 70°N latitude, Eurasian average [3].

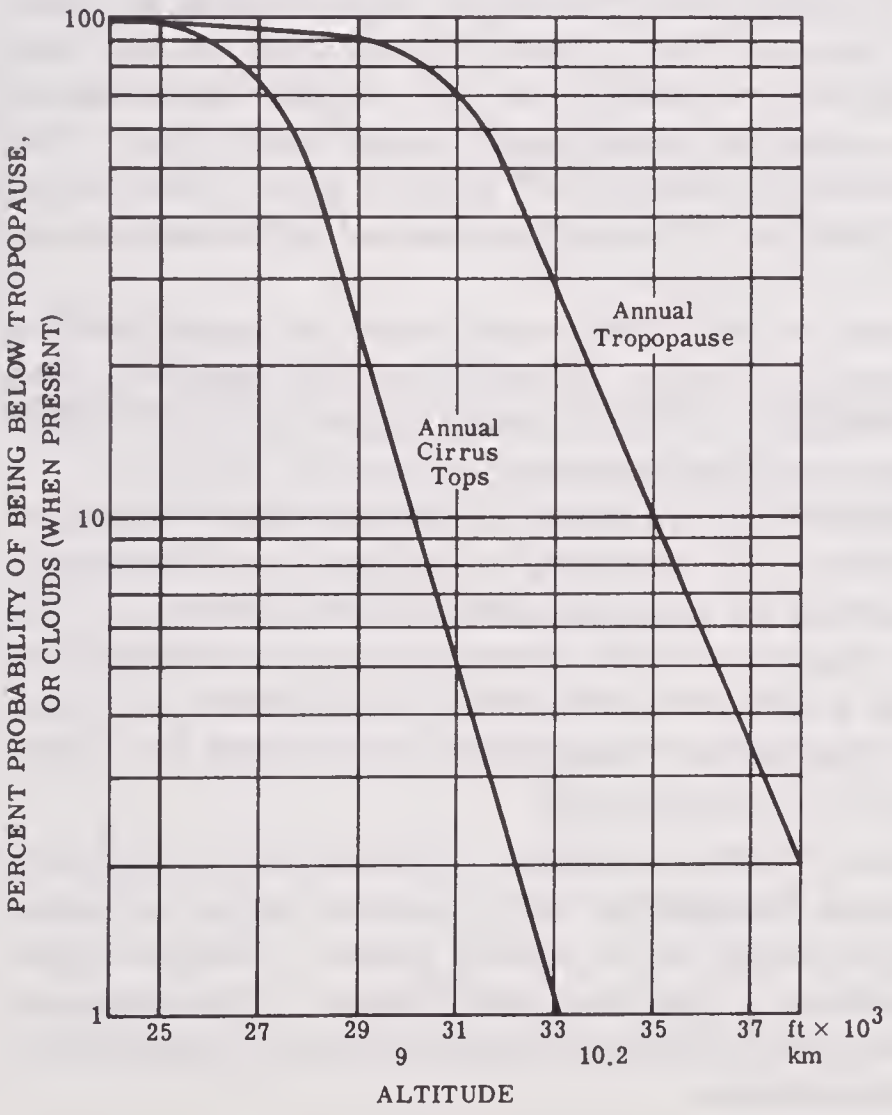


FIG. 5-41. Distribution of tropopause and cloud tops, 65° to 90°N latitude, Eurasian average [3].

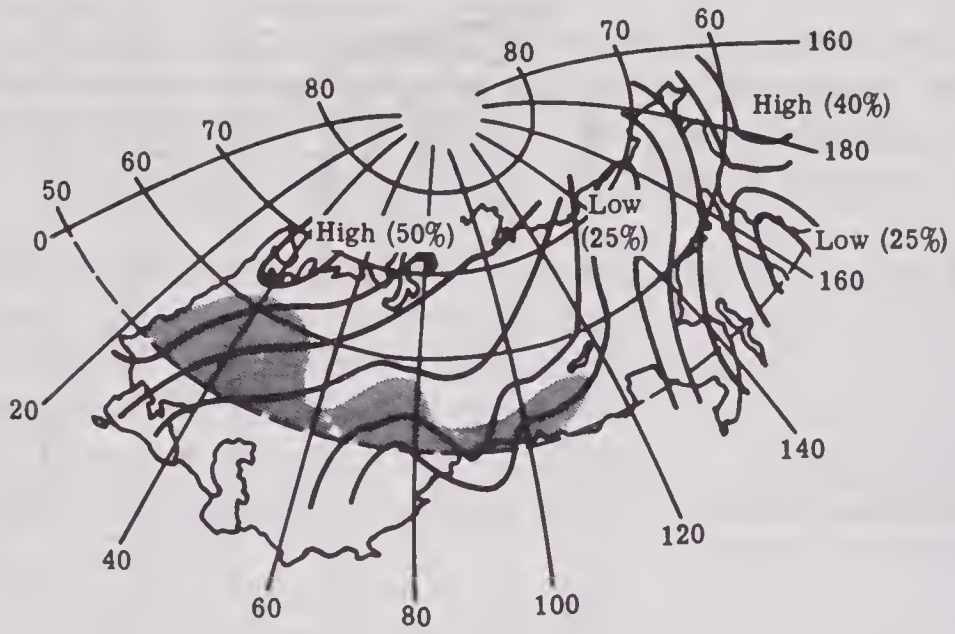


FIG. 5-42. Estimated annual temporal frequency of cirriform clouds. Dotted portion delineates area where 20 or more thunderstorms per year are reported. Overall average = 35% [3].

5.6.2. Stratospheric Clouds [3]. Two types of clouds appear in the upper stratosphere: nacreous clouds at an average height of 24 km, and noctilucent clouds at a height of about 82 km.

Nacreous clouds appear rarely and then mainly in high latitudes characterized by mountainous terrain. They are generally observed in the direction of the sun during sunset or sunrise and are irridescent. Characteristic synoptic conditions that exist with these clouds are strong and consistent northwest winds extending to great heights with below average stratospheric temperatures. Theoretical considerations of water-droplet and ice-crystal growth in nacreous clouds suggest that the radii are less than 1.2 μ, with a very narrow size spectrum of about 0.1 μ. The particle concentration should be essentially that of the available condensation nuclei, about 1/cm^3. The liquid water content would be therefore between 10^{-12} to 10^{-11} g/cm^3. Such liquid water content is lower by about a factor of 10^4 than those observed in the tropospheric clouds.

Noctilucent clouds are visible against the nighttime sky when the upper levels of the atmosphere are still illuminated by sunlight. These clouds have generally been reported only in the Northern Hemisphere during summer (August through October) within a restricted zone of latitudes extending from about 45° to 63°N.

Sunlight scattered from noctilucent clouds exhibits a spectrum and a degree of polarization which can be attributed to the scattering of sunlight by dielectric particles with predominant radii of around 0.1 μ and not greater than 0.2 to 0.4 μ. The observed brightness of the clouds suggests that the corresponding concentrations and matter content should be between 1 and 10^{-2} particles/cm^3 and between 10^{-17} and 10^{-16} g/cm^3, respectively. Such cloud-particle concentrations are about five orders of magnitude less than those given for nacreous clouds.

5.6.3. Probability of Coverage at Various Altitudes. Figures 5-43 through 5-58 are charts showing, for the Northern Hemisphere, altitudes above which the probabilities of less than 0.1 sky coverage are 95, 90, 80, and 60 percent. Charts are presented for the midseason months January, April, July, and October. The criterion of less than 0.1 sky cover (actually less than 0.05 sky cover) can be taken as essentially no interference by clouds for air-to-air operation.

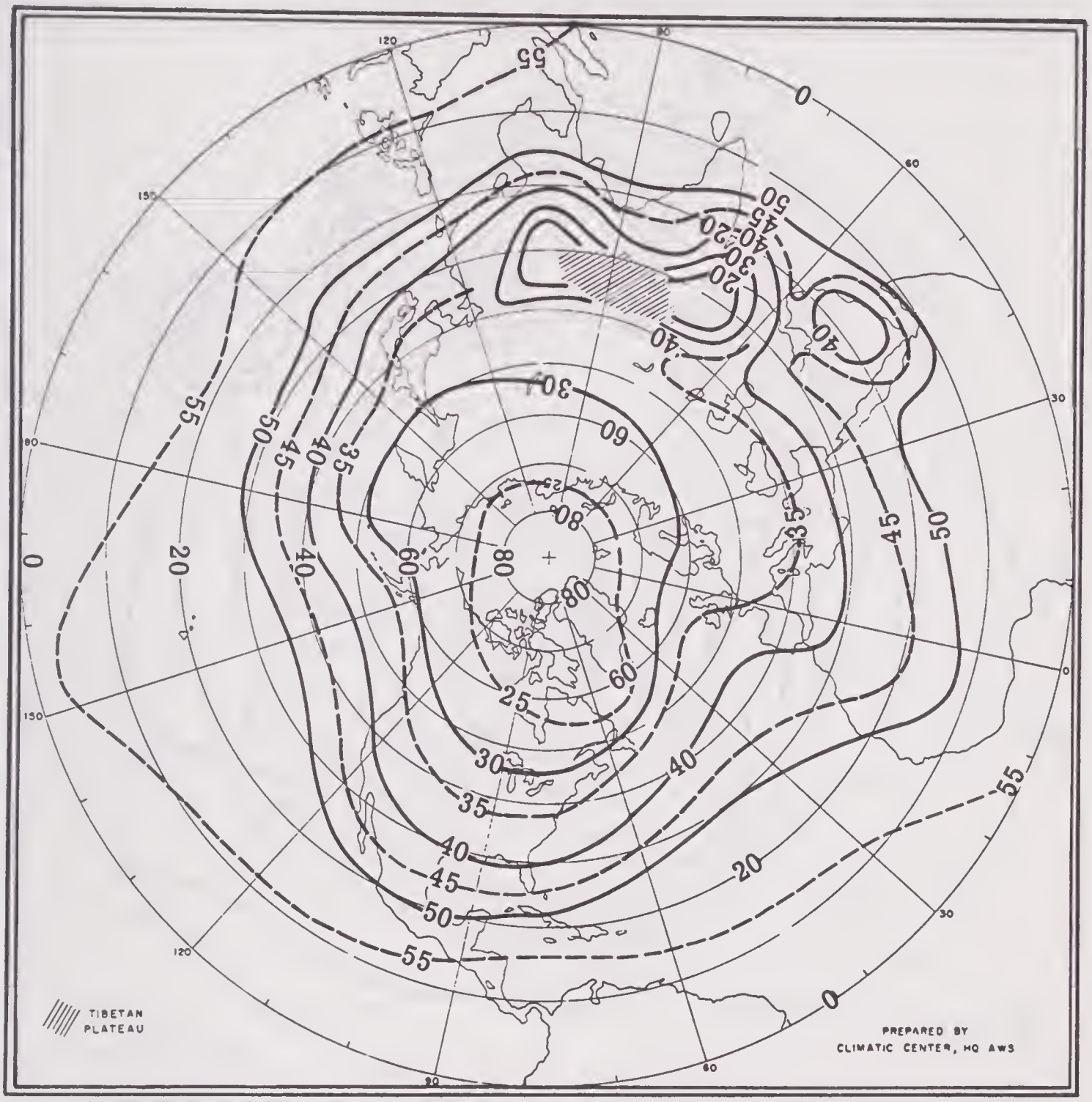


FIG. 5-43. Altitudes (thousands of feet MSL) above which there is 95% probability of having less than 0.1 sky cover, month of January [26].

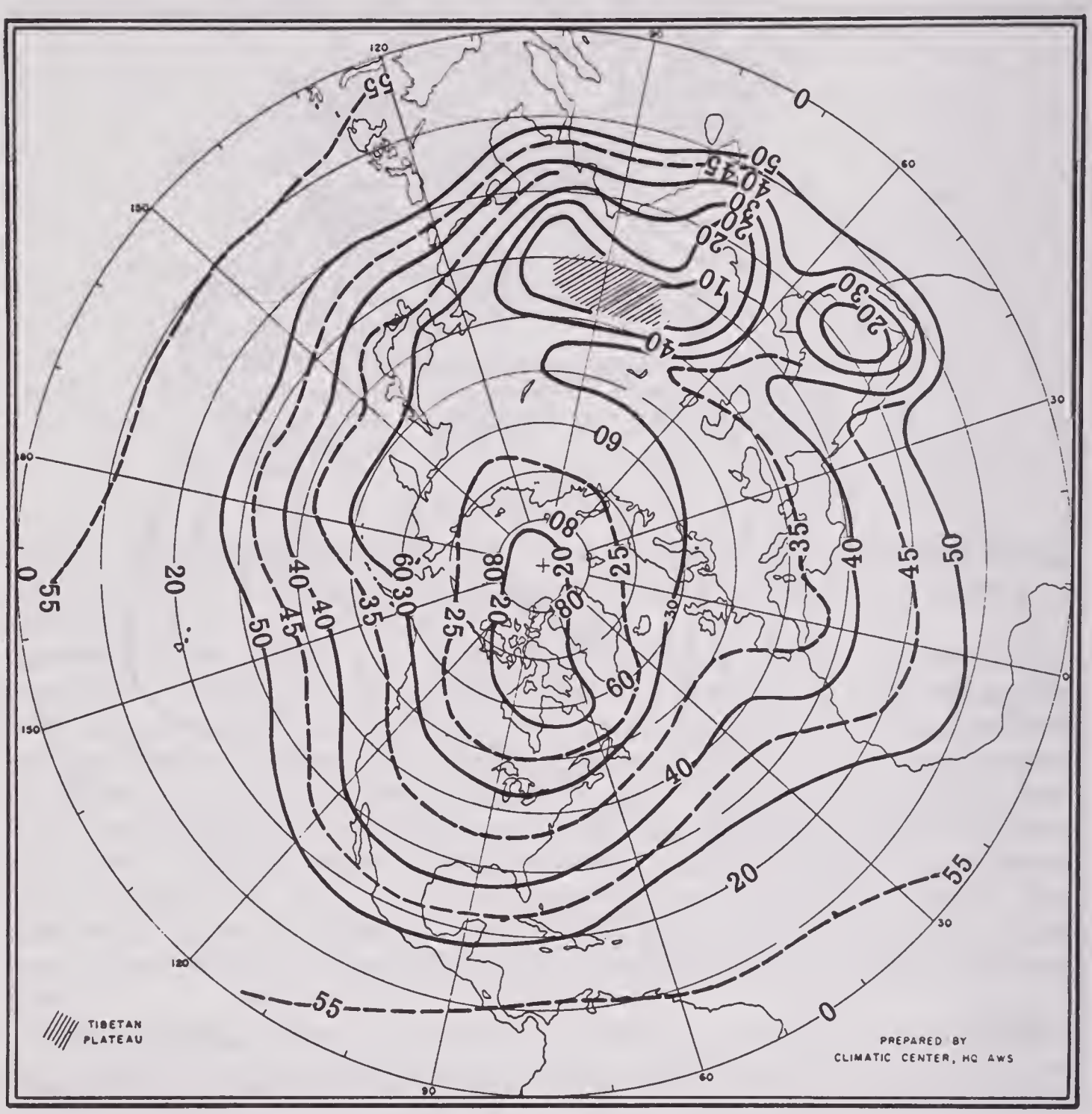


FIG. 5-44. Altitudes (thousands of feet MSL) above which there is 90% probability of having less than 0.1 sky cover, month of January [26].

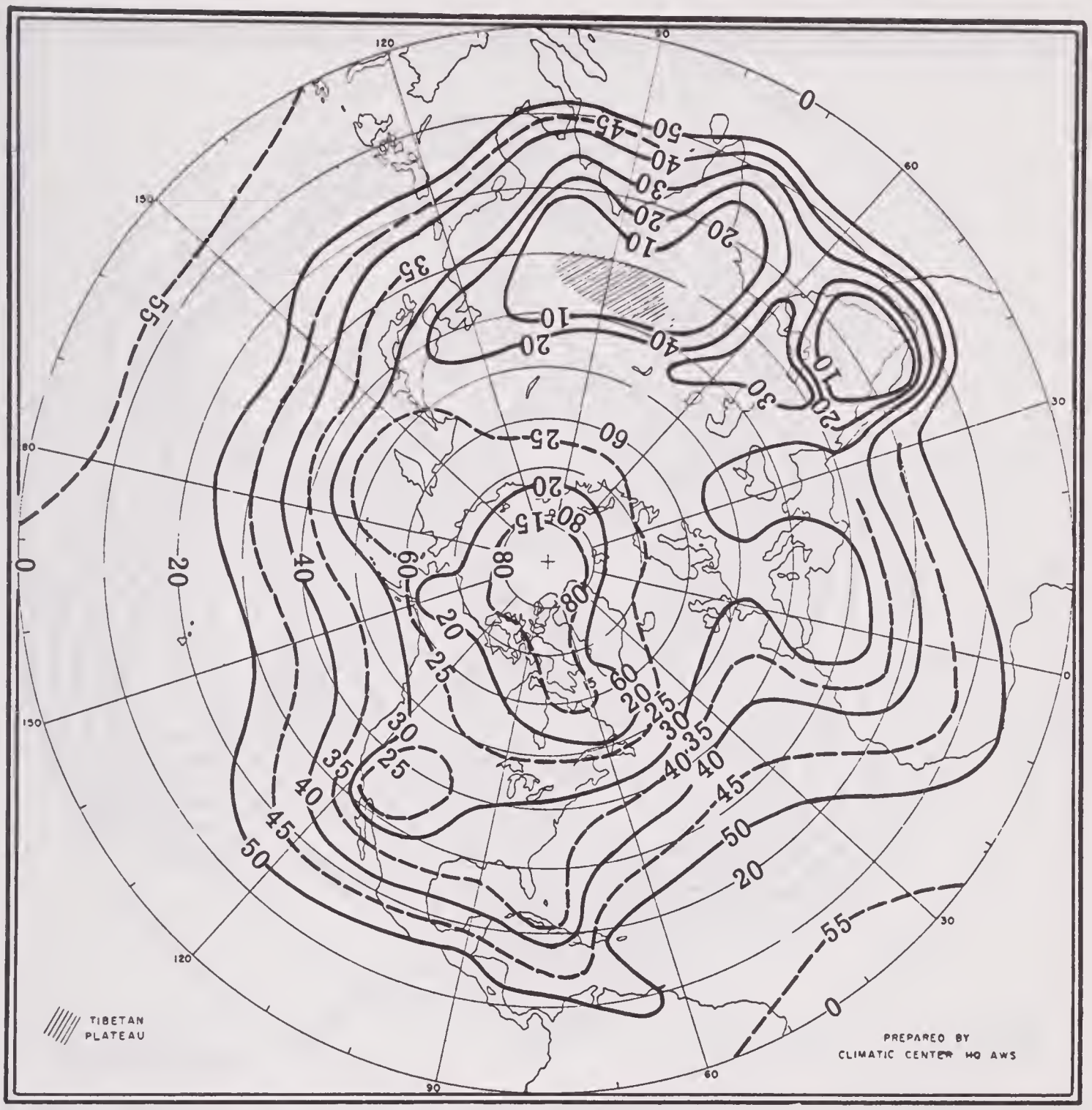


FIG. 5-45. Altitudes (thousands of feet MSL) above which there is 80% probability of having less than 0.1 sky cover, month of January [26].

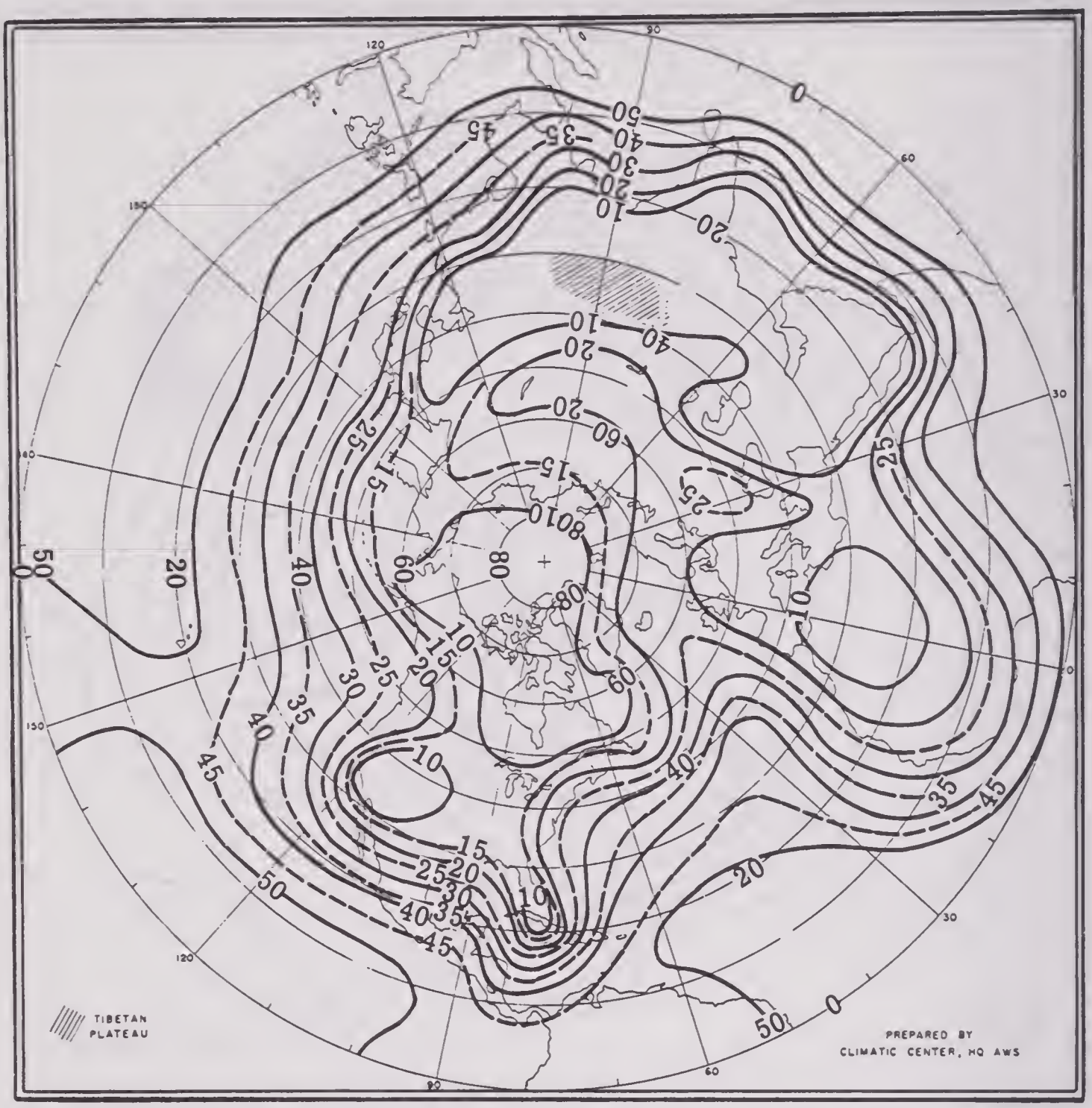


FIG. 5-46. Altitudes (thousands of feet MSL) above which there is 60% probability of having less than 0.1 sky cover, month of January [26].

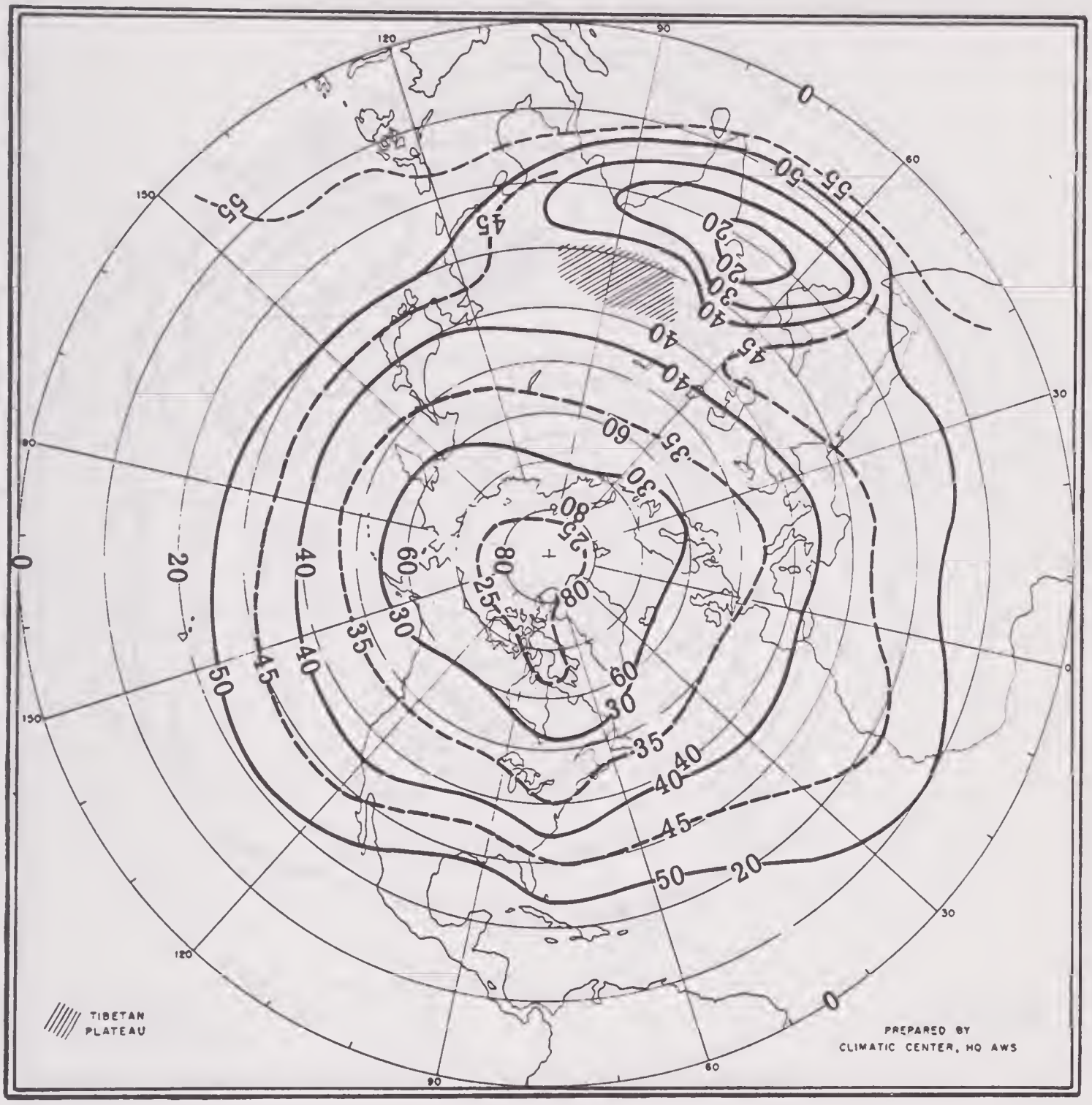


FIG. 5-47. Altitudes (thousands of feet MSL) above which there is 95% probability of having less than 0.1 sky cover, month of April [26].

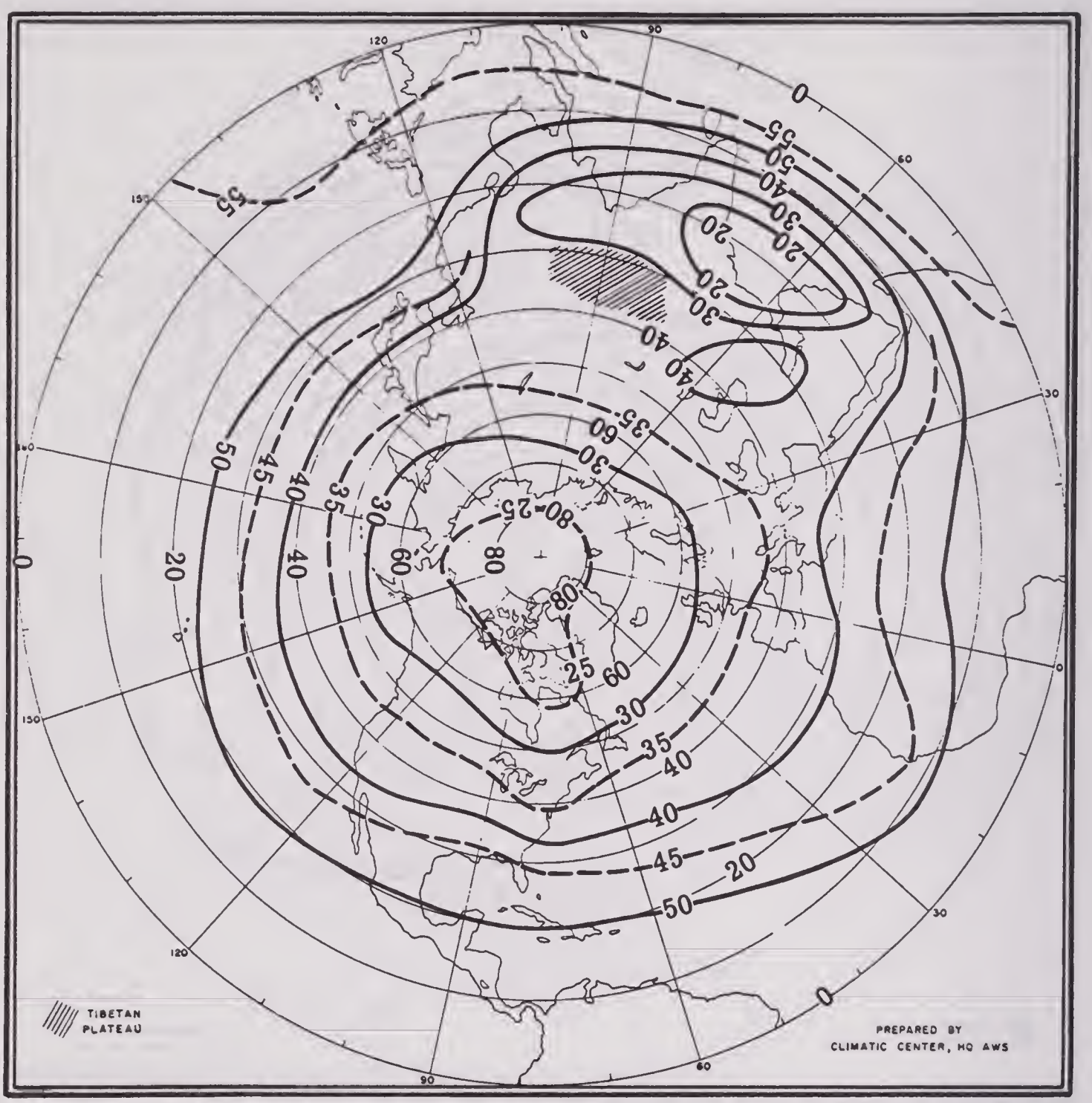


FIG. 5-48. Altitudes (thousands of feet MSL) above which there is 90% probability of having less than 0.1 sky cover, month of April [26].

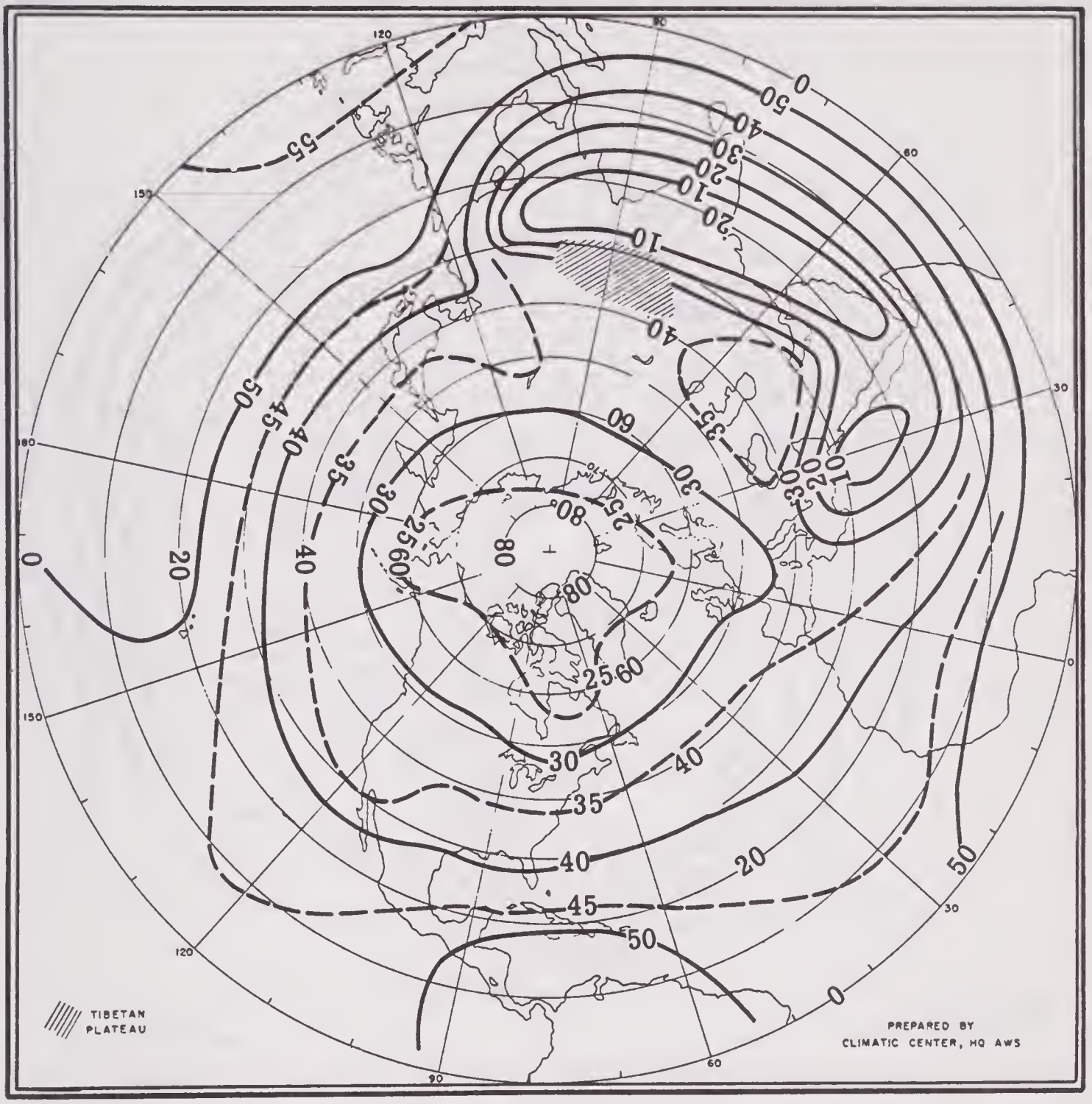


FIG. 5-49. Altitudes (thousands of feet MSL) above which there is 80% probability of having less than 0.1 sky cover, month of April [26].

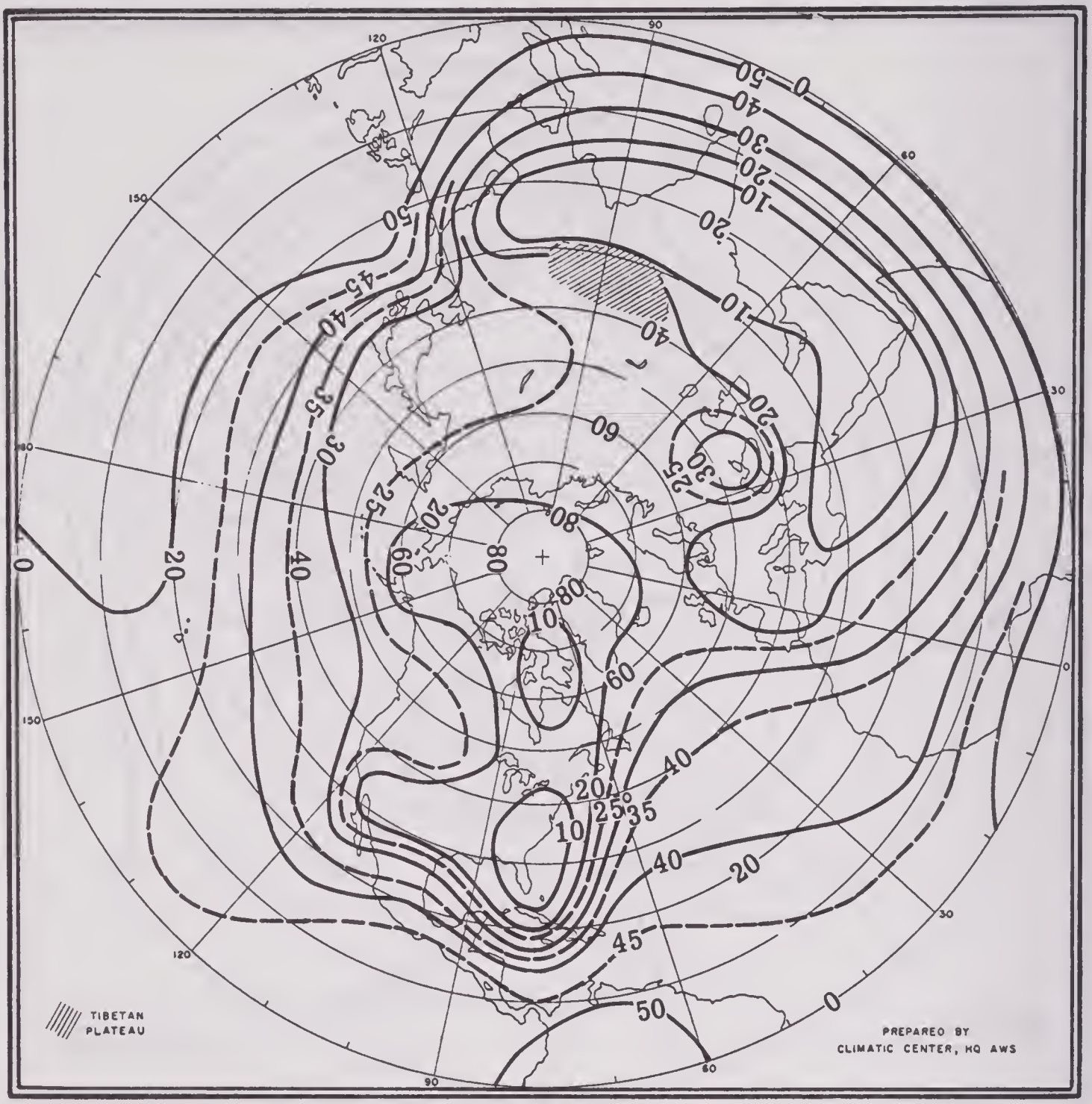


FIG. 5-50. Altitudes (thousands of feet MSL) above which there is 60% probability of having less than 0.1 sky cover, month of April [26].

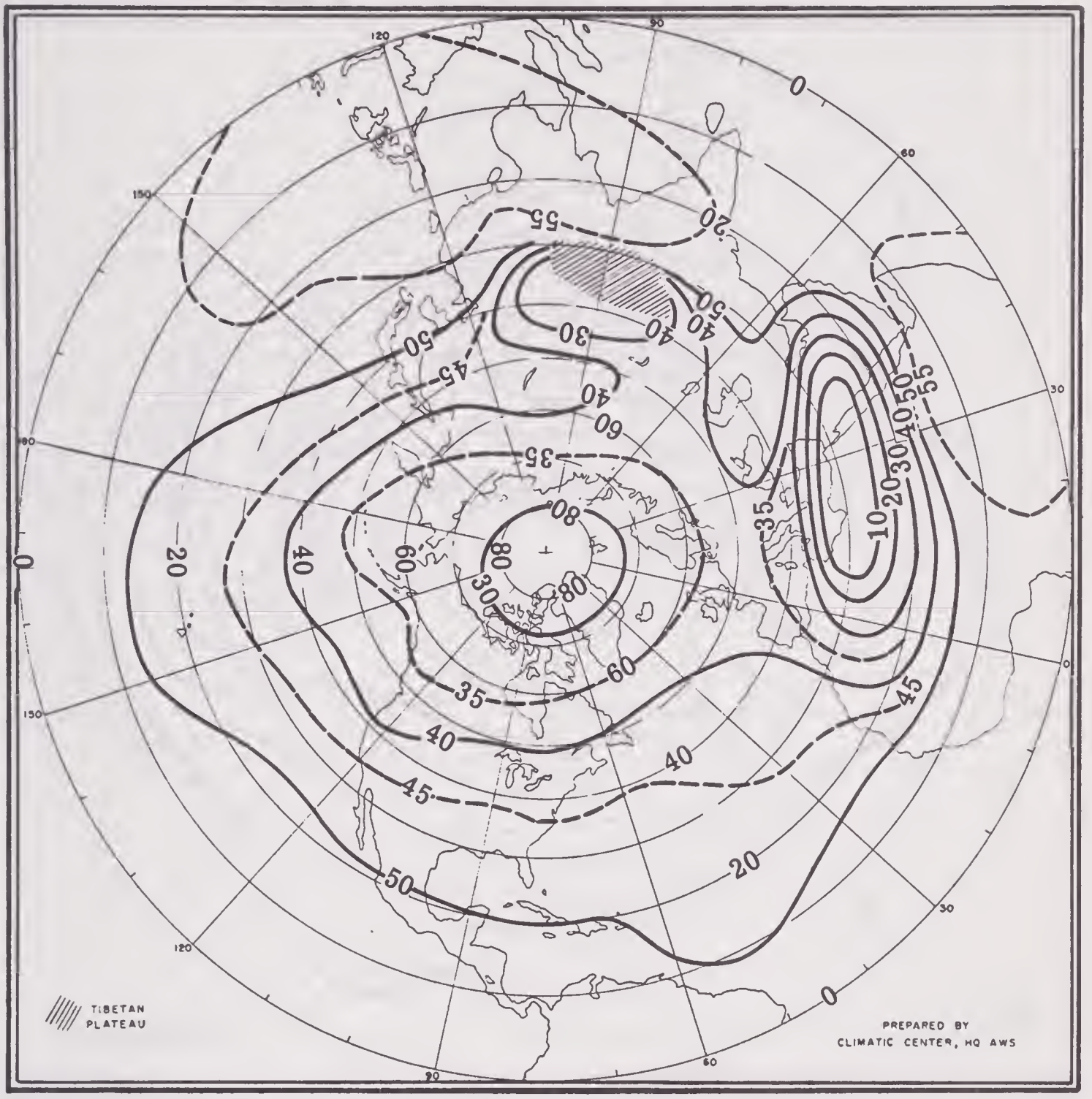


FIG. 5-51. Altitudes (thousands of feet MSL) above which there is 95% probability of having less than 0.1 sky cover, month of July [26].

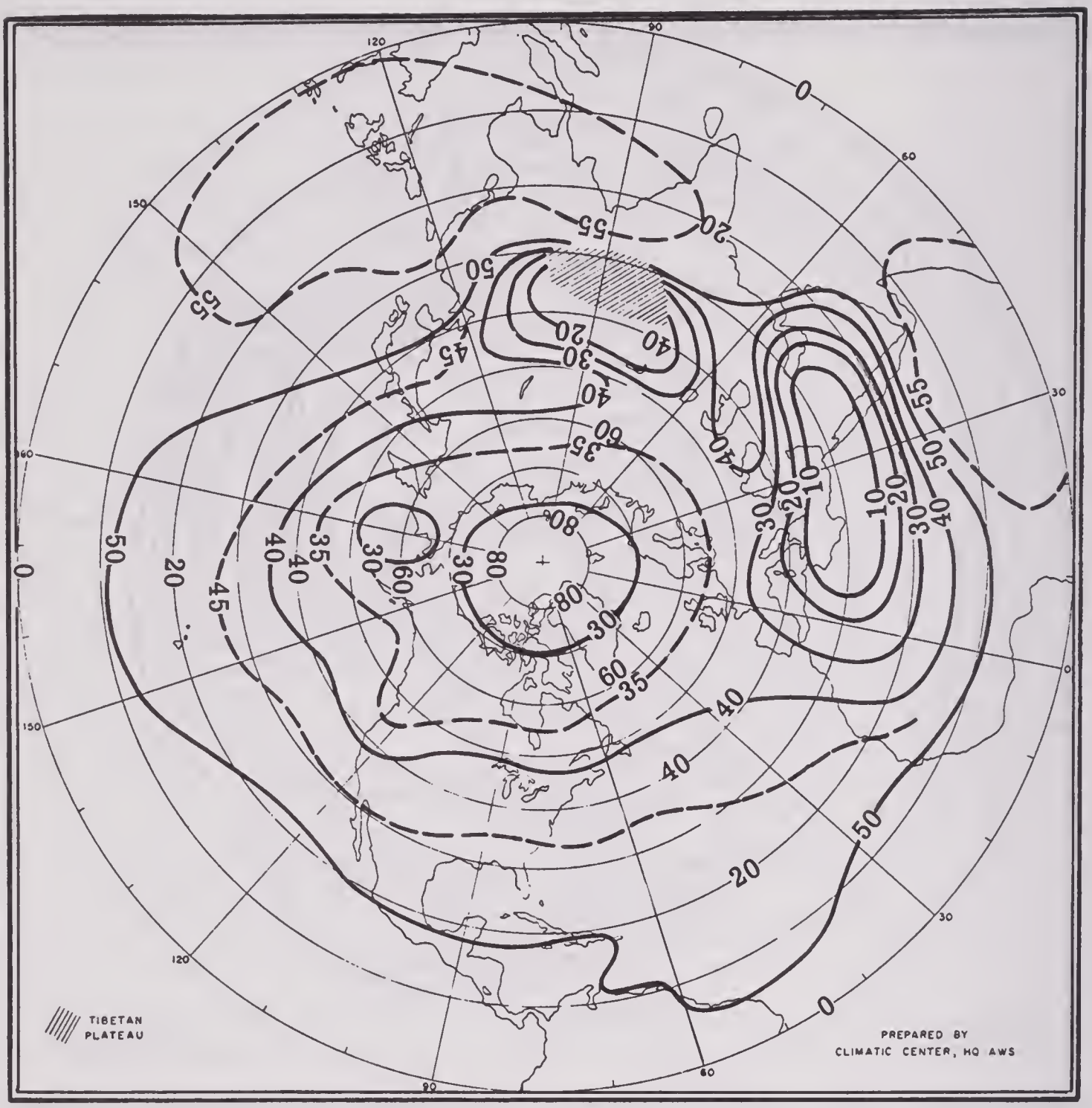


FIG. 5-52. Altitudes (thousands of feet MSL) above which there is 90% probability of having less than 0.1 sky cover, month of July [26].

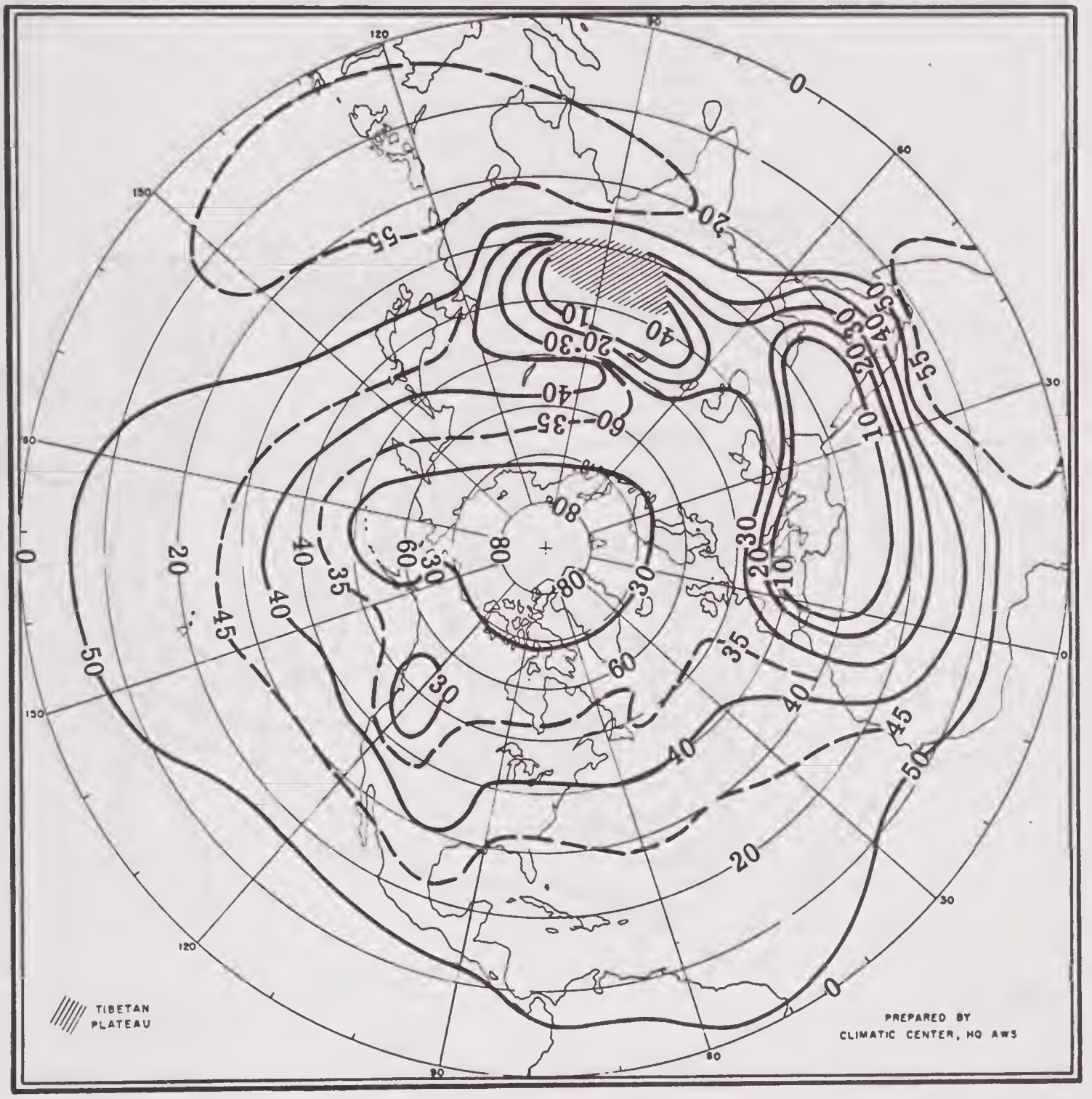


FIG. 5-53. Altitudes (thousands of feet MSL) above which there is 80% probability of having less than 0.1 sky cover, month of July [26].

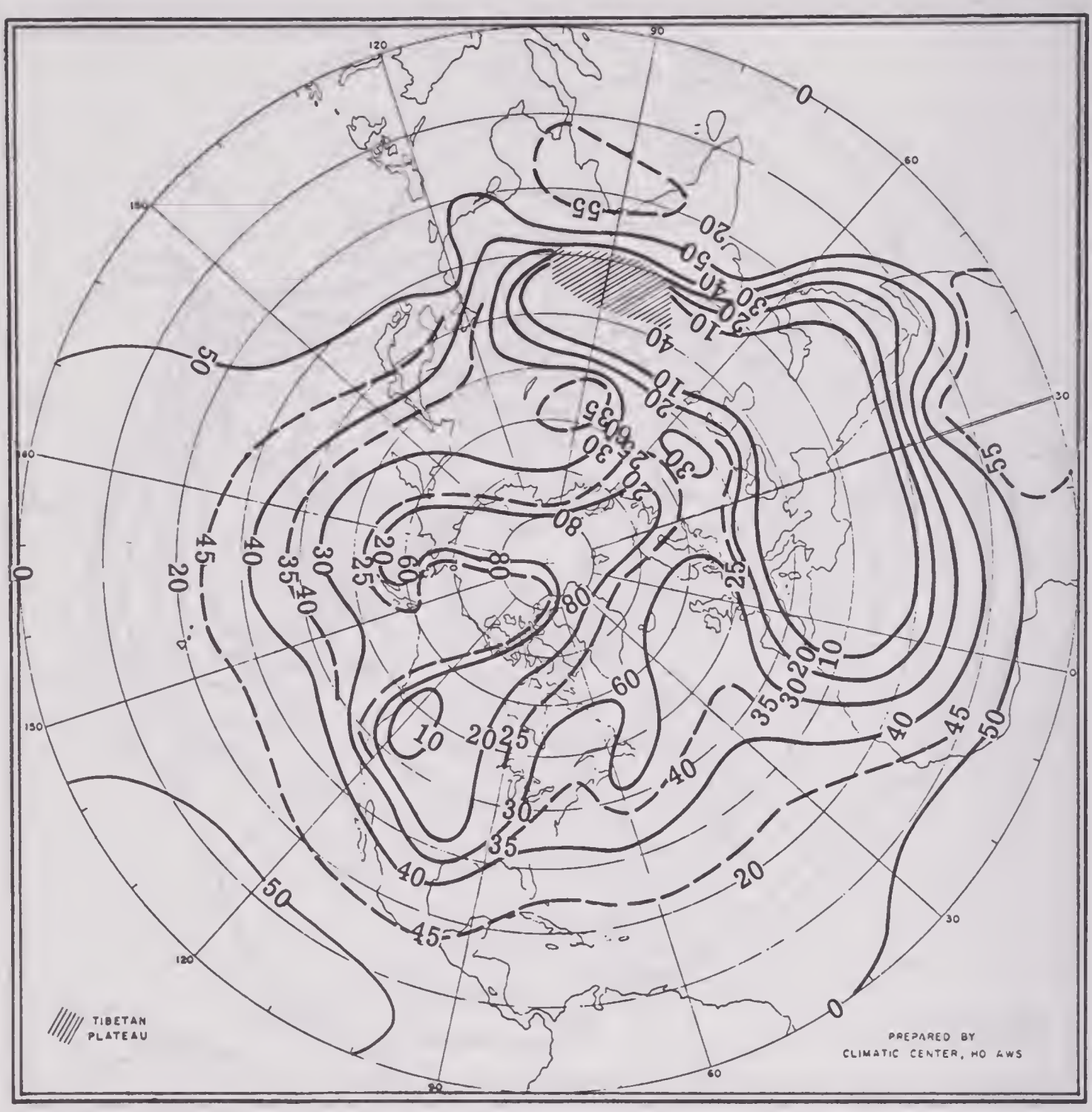


FIG. 5-54. Altitudes (thousands of feet MSL) above which there is 60% probability of having less than 0.1 sky cover, month of July [26].

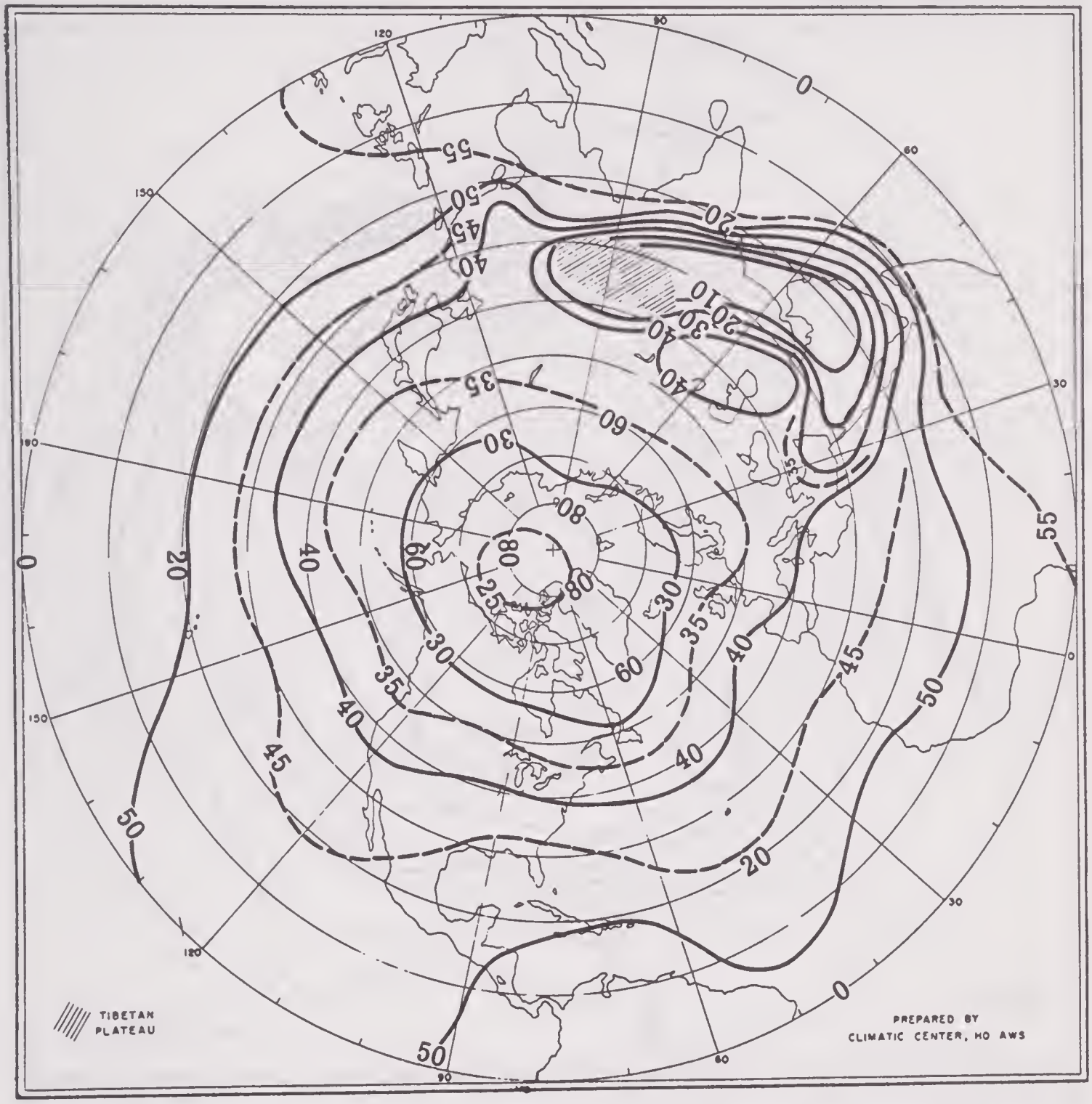


FIG. 5-55. Altitudes (thousands of feet MSL) above which there is 95% probability of having less than 0.1 sky cover, month of October [26].

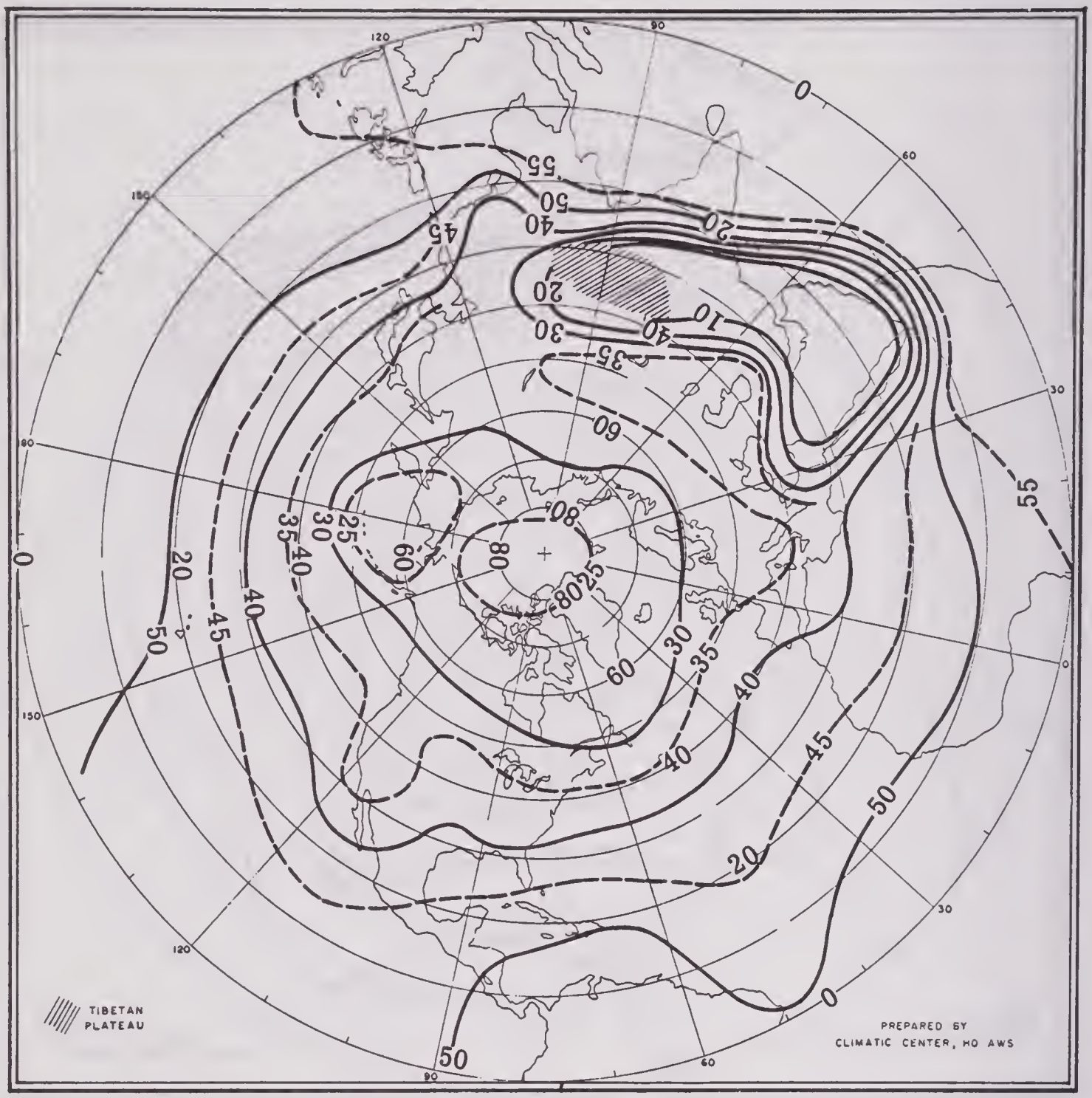


FIG. 5-56. Altitudes (thousands of feet MSL) above which there is 90% probability of having less than 0.1 sky cover, month of October [26].

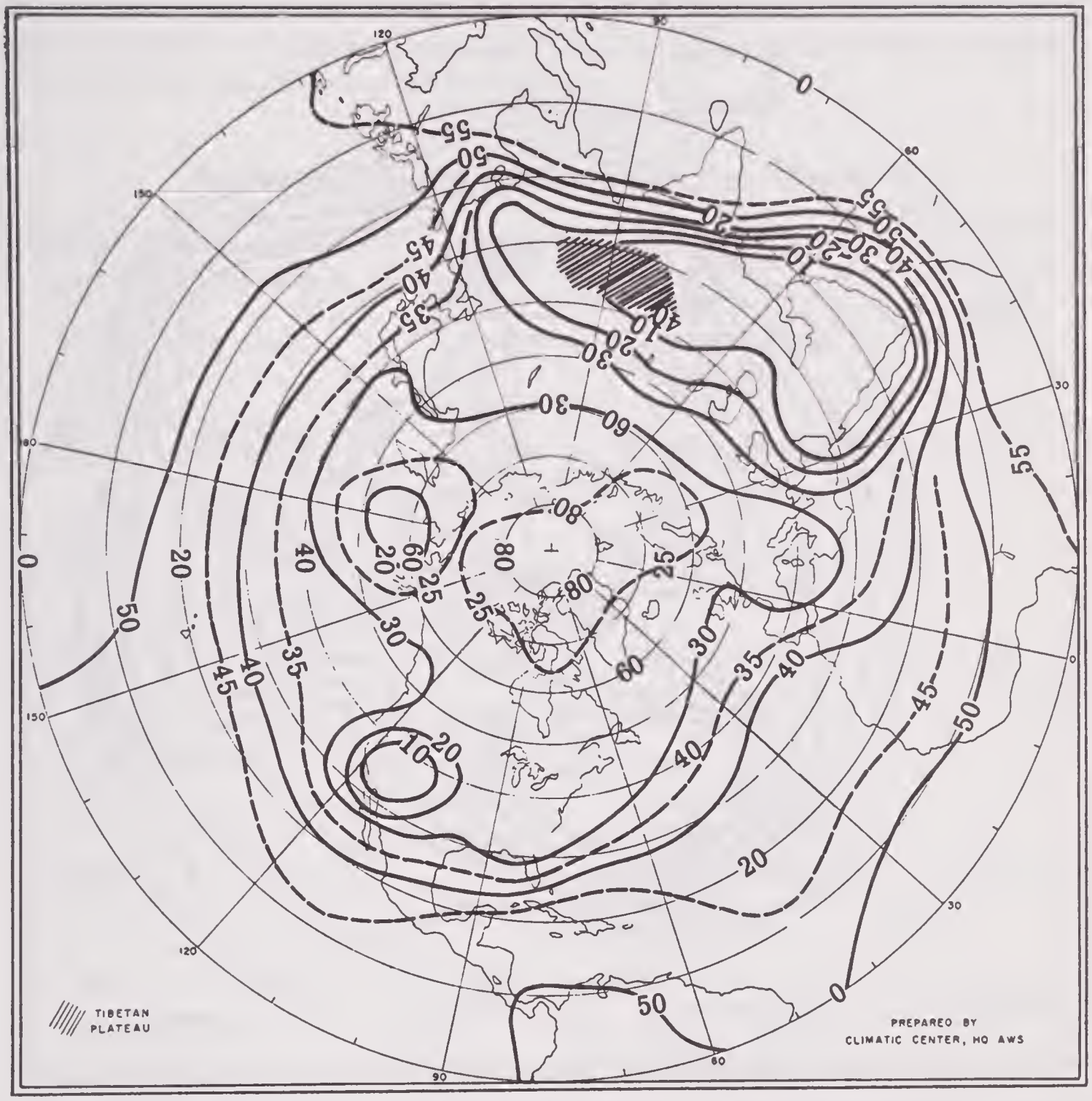


FIG. 5-57. Altitudes (thousands of feet MSL) above which there is 90% probability of having less than 0.1 sky cover, month of October [26].

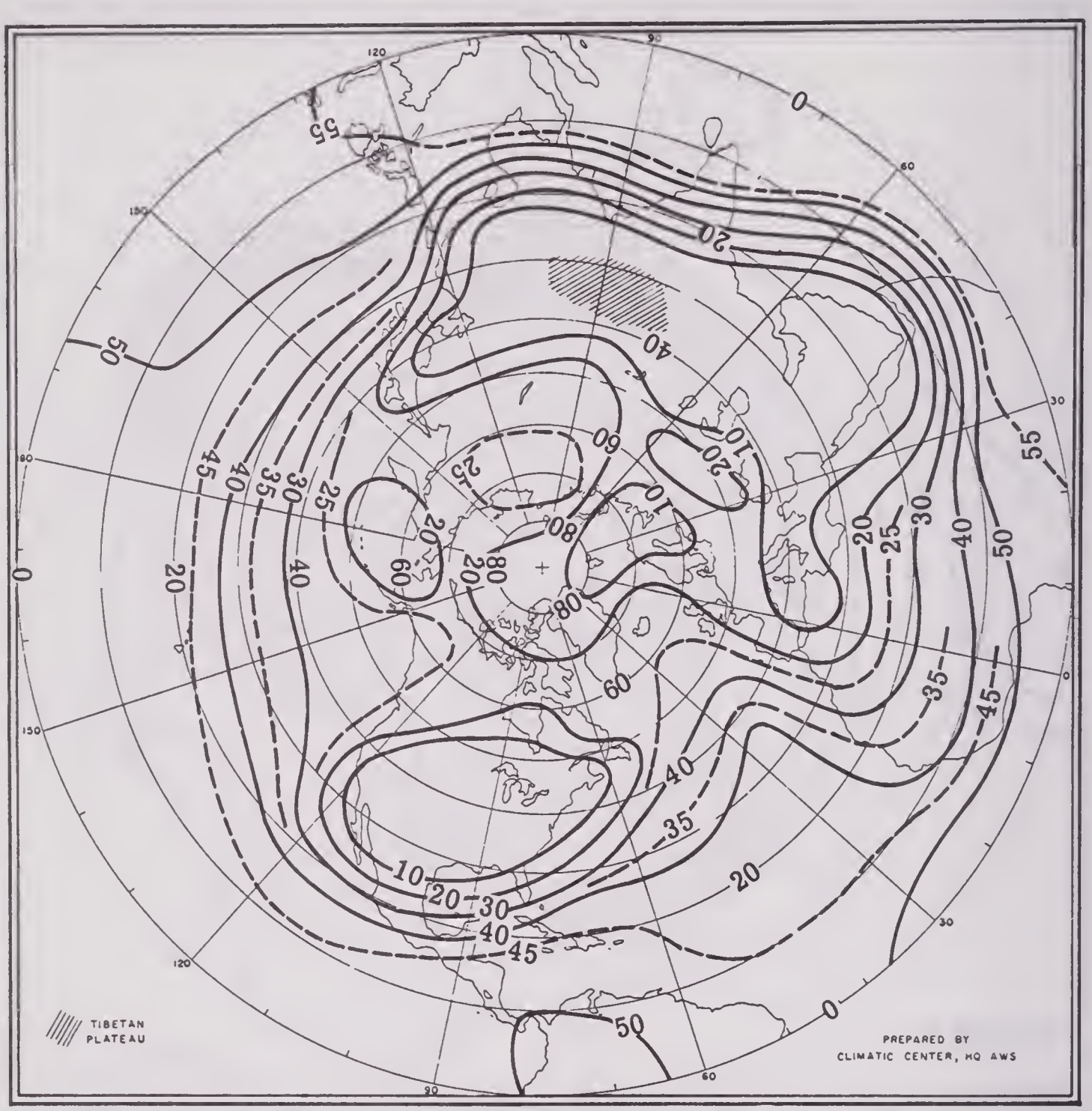


FIG. 5-58. Altitudes (thousands of feet MSL) above which there is 60% probability of having less than 0.1 sky cover, month of October [26].

5.7. Stratospheric Aerosols [3, 63-68]

A uniform distribution of the stratospheric aerosol content tends to decrease infrared gradients. Apparently the stratosphere will contribute to background noise because of the tendency of particles to form clouds which become arranged in periodic structures. Intensities will be high at small scattering angles, and atmospheric attenuation is negligible at these high altitudes. Table 5-10 presents a summary of information on the particle content of the stratosphere.

TABLE 5-10. PARTICLE CONTENT OF THE STRATOSPHERE [3]

Altitude (km)	*Concentrations* (no. per cm^3)	*Radii* (μ)	*Remarks*	*Typical Band Spacings* (km)
10-30	10^{-1} to 1	>0.08 or ~0.10	Stable dust layer, 17 to 22 km –sulfur	–
10-30	10^{-2} to 10^{-1}	~0.15	Stable dust layer, 17 to 22 km –sulfur	–
10-30	~<10^3	~0.8 (horizontal orientation)	Temporary layers of volcanic pumice	1
17-31	~<1	~1.5	Nacreous clouds consisting of ice crystals	40
30-80	1	~0.1 (assumed)	Theoretical by measurements of conductivity	–
74-92	10^{-2}	~0.1	Noctilucent cloud (dust layer or ice crystal)	10 and 60
80	10^{-4} to 10^{-1}	0.1 (assumed)	Theoretical interplanetary dust sources, 10^{-23} to 10^{-20} g/cm^3	–

Stratospheric particulate matter may be divided into two classes: dust particles and condensed water.

Catastrophic volcanic eruptions and forest fires have deposited vast quantities of dust in the upper atmosphere; these can indirectly increase the upward intensity of reflected sunlight by acting as nucleating agents for ice. Unusual concentrations of clouds might result.

The earth is surrounded by belts of dust, smoke, and ice particles. Encounters with the dust by the earth's gravitational field causes an accretion of 10 to 50 lb of matter per square mile per year, based on an estimate of 24,000,000 visible meteors per day [25]. Some of this dust is concentrated into two extreme outer shells: the lighter smoke between altitudes of 2000 to 4000 mi, and dust from 600 to 1000 mi.

In general, the total particle concentrations just above the tropopause are between 10 and 100/cm^3, but decrease to 1/cm^3 or less above about 20 km.

The manner in which a particle scatters light depends on the ratio of its radius to the wavelength of light. For ratios up to about 0.08, Rayleigh's laws hold; between 0.08 and 3 the Mie theory is used; and at larger values, the laws of geometrical optics are satisfactory. Figure 5-59 gives examples of particle sizes. Figure 5-60 shows the concentration of different particle sizes at various altitudes.

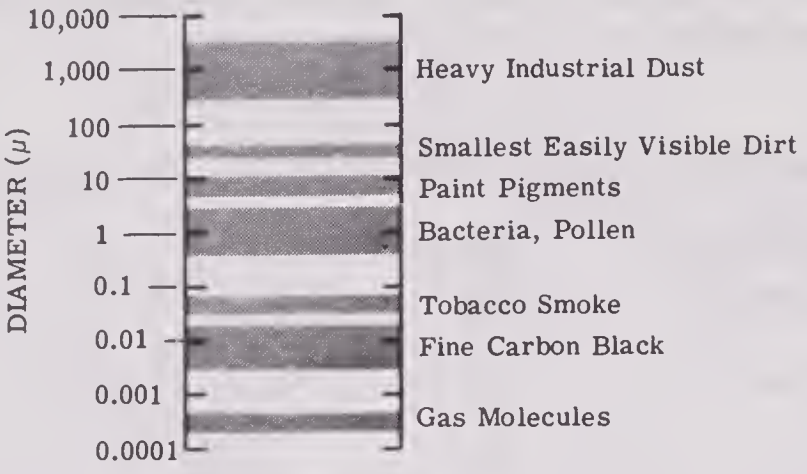


FIG. 5-59. Examples of particle sizes [3].

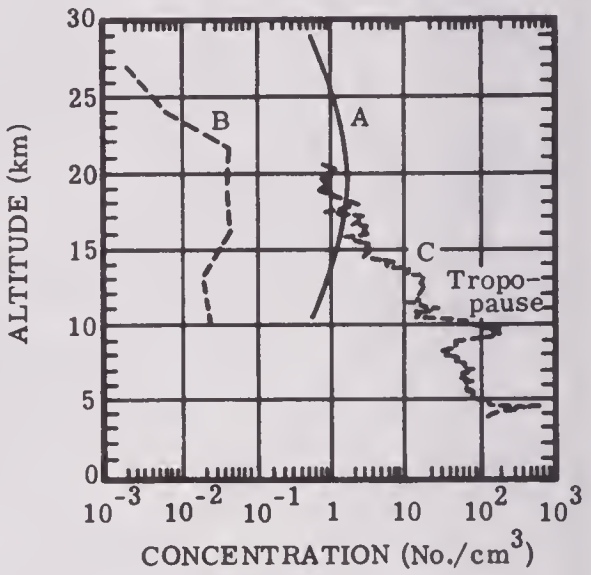


FIG. 5-60. Vertical profiles of particle concentrations. Curve A = radii greater than 0.08 μ; curve B = radii greater than 0.1-0.3 μ; curve C = Aitken nuclei (0.01- to 0.1-μ radii) [3].

5.8. Spectral Radiance of Terrain

The apparent spectral radiances discussed in this section include thermal emission and radiation reflected by the ground, as well as scattered and emitted radiation contributed by the atmosphere in the line of sight.

5.8.1. Terrain Temperature. Soil or other terrestrial surfaces have a mean temperature value of approximately 300°K, and peak radiance is near 10 μ. The earth's surface temperature depends upon the incident solar radiation and the radiative boundary conditions as well as conductive and convective processes. The latter processes depend on the physical and chemical characteristics of particular components of the terrain and the local weather conditions.

5.8.2. Terrain Emissivity and Reflectivity.* The amount of radiation that is absorbed, reflected, or scattered varies with wavelength and with the nature of the terrain.

The reflectance values for natural objects, at wavelengths shorter than 3 μ, range from 0.03 for bare ground or ocean to 0.95 for fresh snow [1,11,24,27,28,29]. In the long-wavelength region reflectance values range from practically zero to 0.72 [30-32].

*See also *Notes Added in Proof*, page 171.

5.8.3. Spectra in the Emission Region. Figure 5-61 shows the comparative spectral radiances of a patch of ground at an airfield (Colorado) observed on a clear night, and during the following morning with the sun shining on it.

Figure 5-62 shows the radiance of the night sky just above the horizon and that of the ground at the same angle below the horizon.

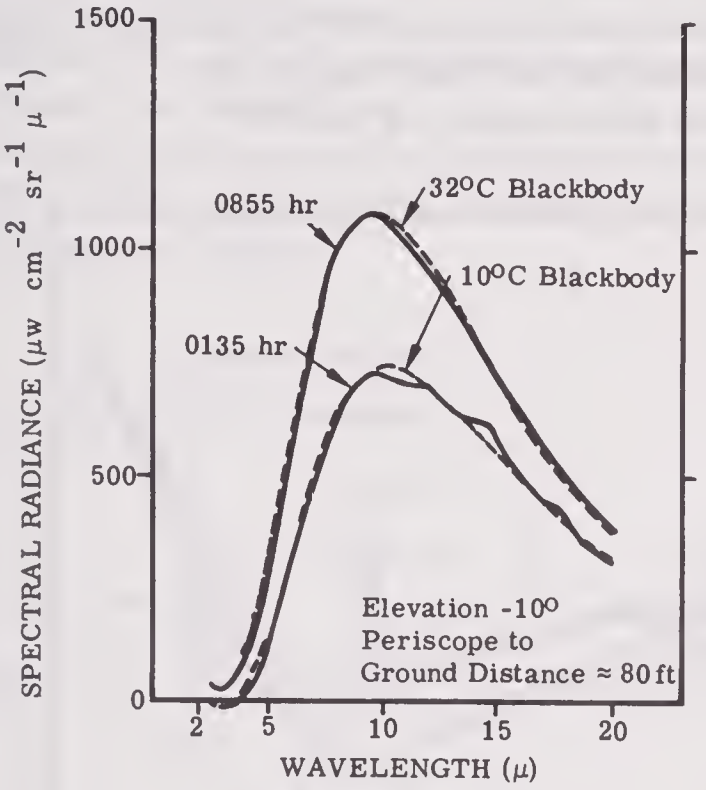


FIG. 5-61. Day and night radiances of grass-covered field (Peterson Field, Colorado) [33].

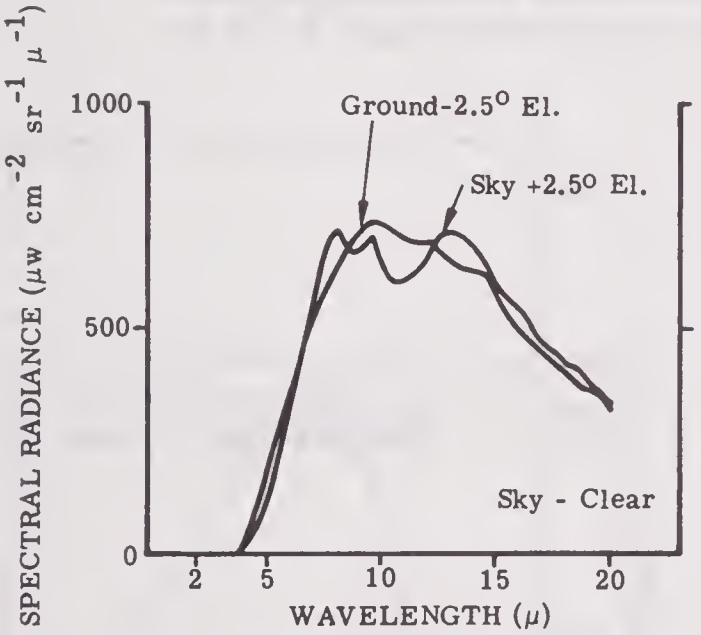


FIG. 5-62. Comparative spectra of the ground and sky near the horizon (Peterson Field, Colorado) [33].

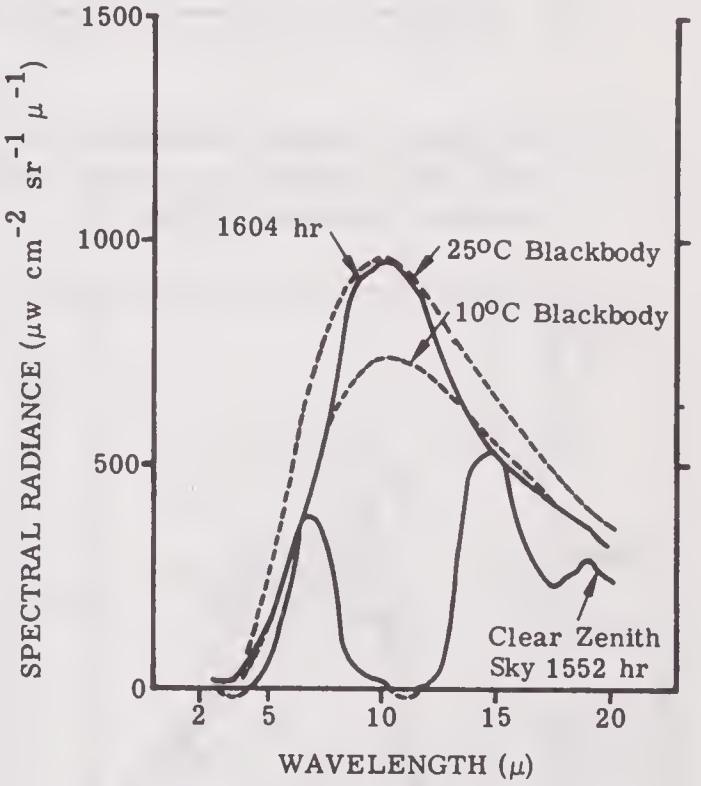


FIG. 5-63. Radiance of an urban area and of clear zenith sky (Colorado Springs from Pikes Peak) [33].

The spectra of distant terrain do not always conform to the blackbody characteristics observed in the radiance of nearby terrain. This can be seen in Fig. 5-63, where the upper curve represents the radiance of a city (Colorado Springs) on a plain as viewed from the summit of a mountain (Pikes Peak) at a distance of about 15 mi. The situation illustrated in Fig. 5-63 is, in a sense, the reverse of that shown in Fig. 5-7 where a cooler cloud was seen from a lower and warmer environment.

Figure 5-64 shows the diurnal variation in the 10-μ radiance of selected backgrounds on the plains, measured from the summit of a mountain (Pikes Peak). The line-of-sight distances are: forest, 30 mi; grassy plains, 21 mi; airfield, 19 mi; city, 15 mi.

Figure 5-65 shows the spectral radiance of dry sand (Cocoa Beach, Florida). The 9-μ dips in the reflectance of the sand for curves *A* and *C* correspond with a wavelength of relative poor emissivity. The reason is that the crystals of common silica sand exhibit reststrahlen at 9 μ. With overcast sky (curve *B*), the added sky radiance reflected at this wavelength just compensates for the loss of emissivity.

The effect of moisture on the radiance of sand is seen in Fig. 5-66.

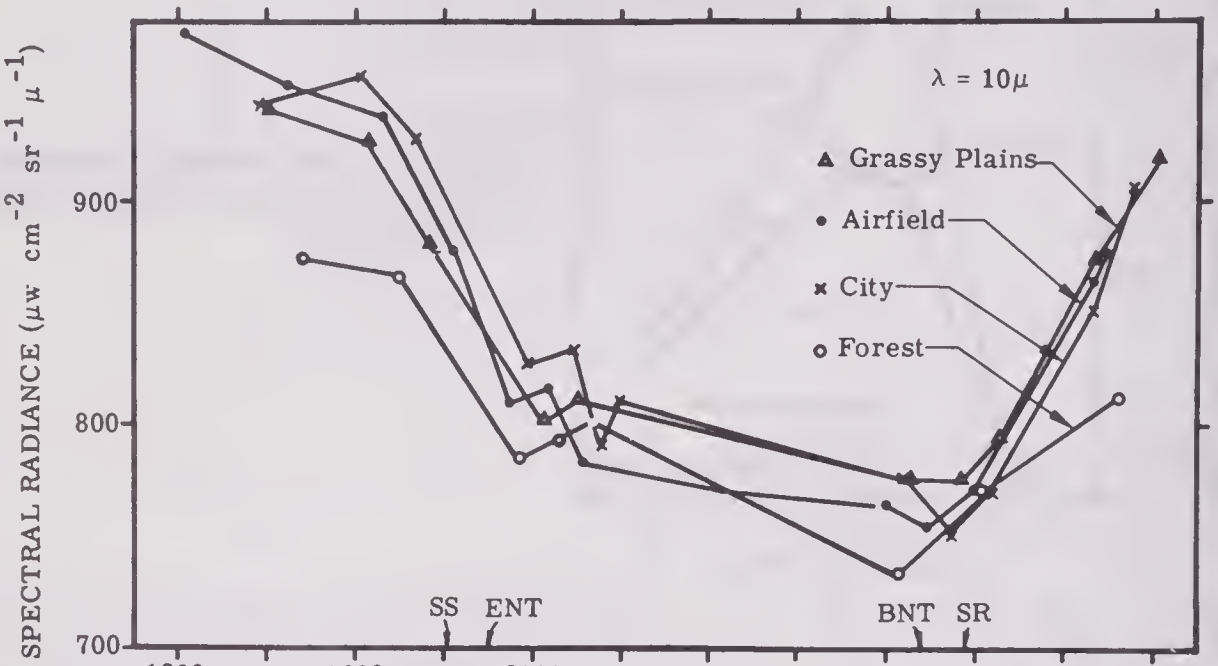


FIG. 5-64. Diurnal variation in the 10-μ radiance of selected backgrounds [33]. SS = sunset; SR = sunrise; ENT = end of nautical twilight; BNT = beginning of nautical twilight.

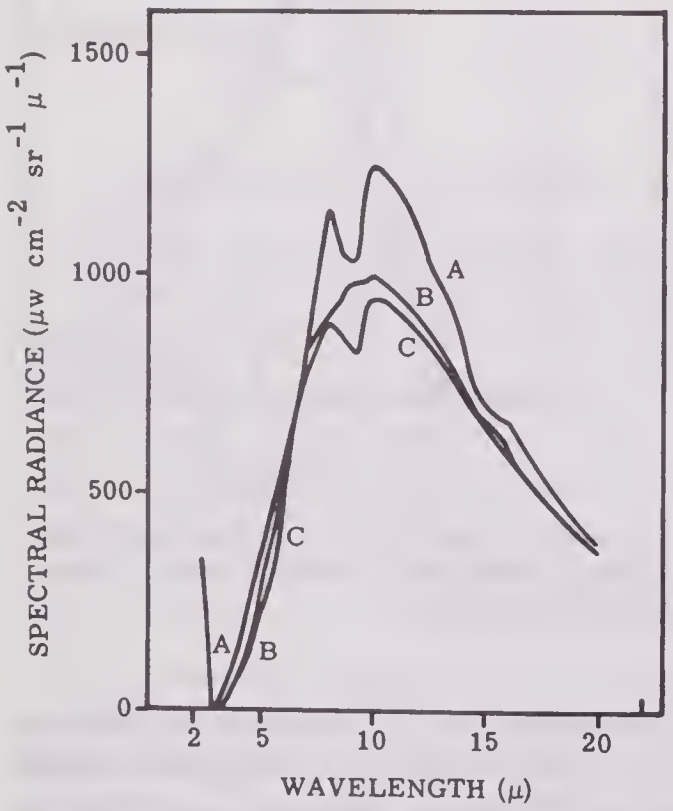


FIG. 5-65. Spectral radiance of dry sand (Cocoa Beach, Florida) [33]. *A* = sunlit sand, *B* = sand under a cloudy night sky, *C* = sand on a clear night.

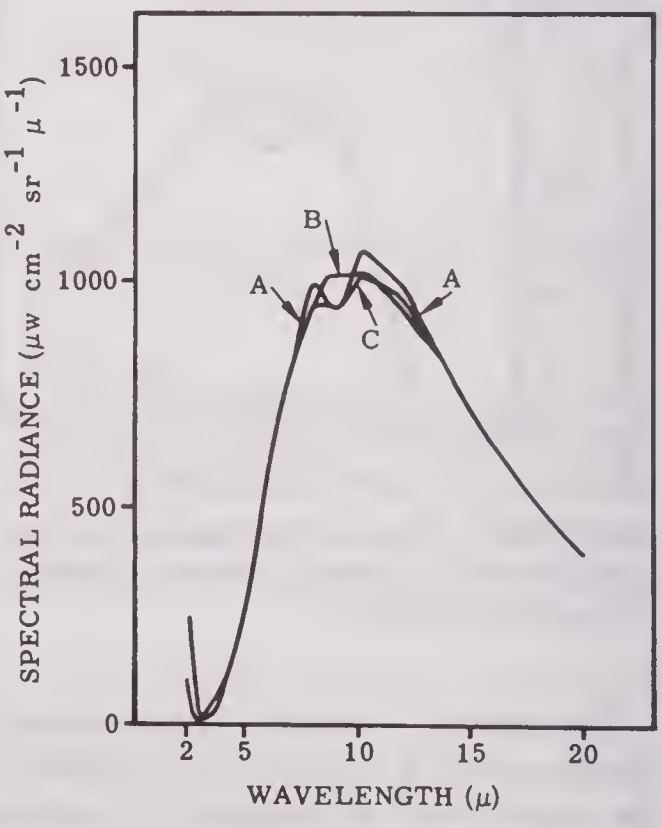


FIG. 5-66. Spectral radiance of moist sand, (Cocoa Beach, Florida); *A* = dry sand, *B* = extremely wet sand, *C* = moist sand [33].

5.8.4. Spectra in the Scattering Region. The daytime spectra of objects at ambient temperatures show minima around 3 to 4 μ. In the more transparent regions of the spectrum between 3 and 5 μ, in the daytime, the sky a few degrees above the horizon radiates less than the ground a few degrees below. In the scattering region of the spectrum, the sky and the ground often show radiances of comparable magnitude. Usually the ground radiance is somewhat higher than that of the sky, and is frequently a minimum near the horizon.

Figures 5-67 and 5-68 are elevation scans of the spectral radiance near the horizon at different wavelengths.

The scan covers alternating patches of shaded and sunlit ground, trees, mountains (northern slope of Pikes Peak, Colorado), and the clear sky up to approximately 15° or 20° near the horizon.

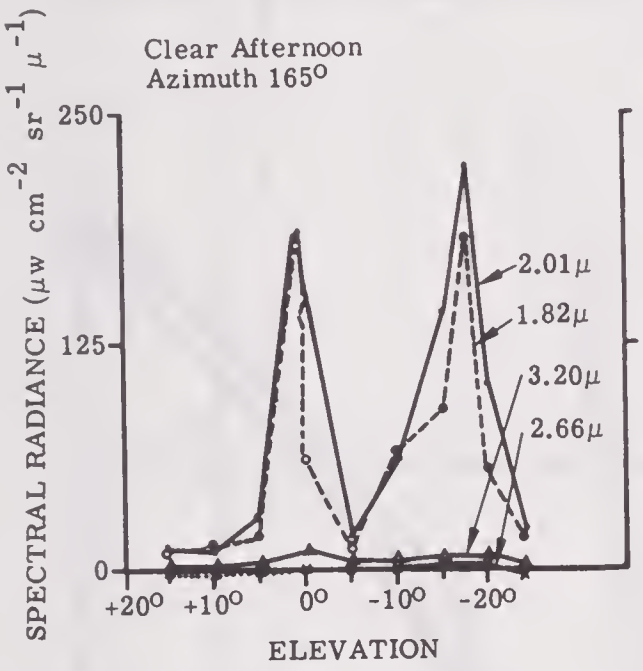


FIG. 5-67. Elevation scans across mountainous terrain at fixed wavelengths in the interval 1.8 to 3.2 μ [33].

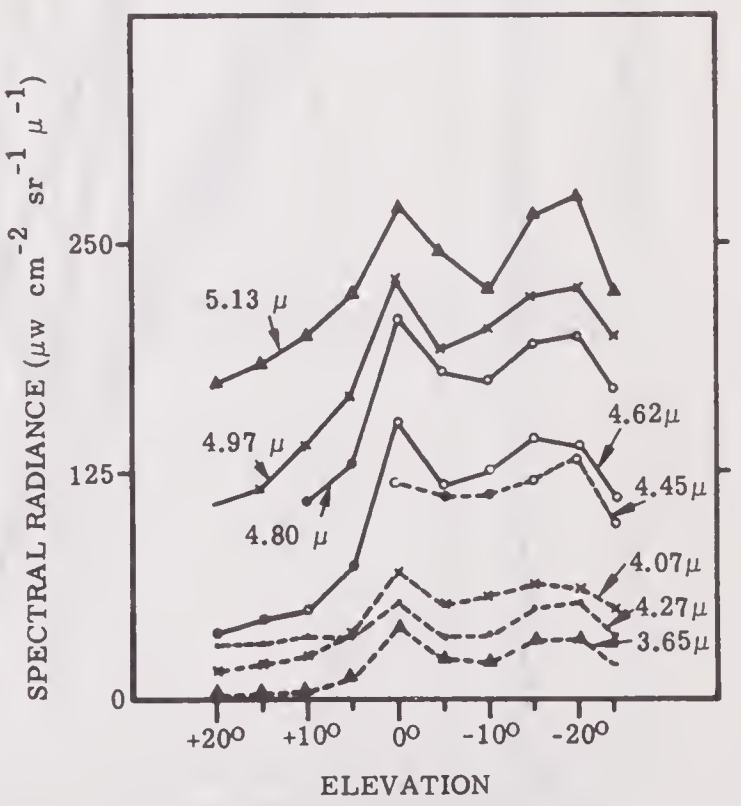


FIG. 5-68. Elevation scans across terrain at the same time as in Fig. 5-67 at fixed wavelengths in the intervals 3.5 to 5.2 μ [33].

5.8.5. Spectral Radiance of Various Objects and Surfaces [31,34,35]. Figures 5-69 through 5-88 are measurements of the infrared spectral radiance from various objects and surfaces. These measurements were made under different temperatures, humidity, sky conditions, etc. In the figures, α_H and α_V are the angular fields of view in the horizontal and vertical directions in object space.

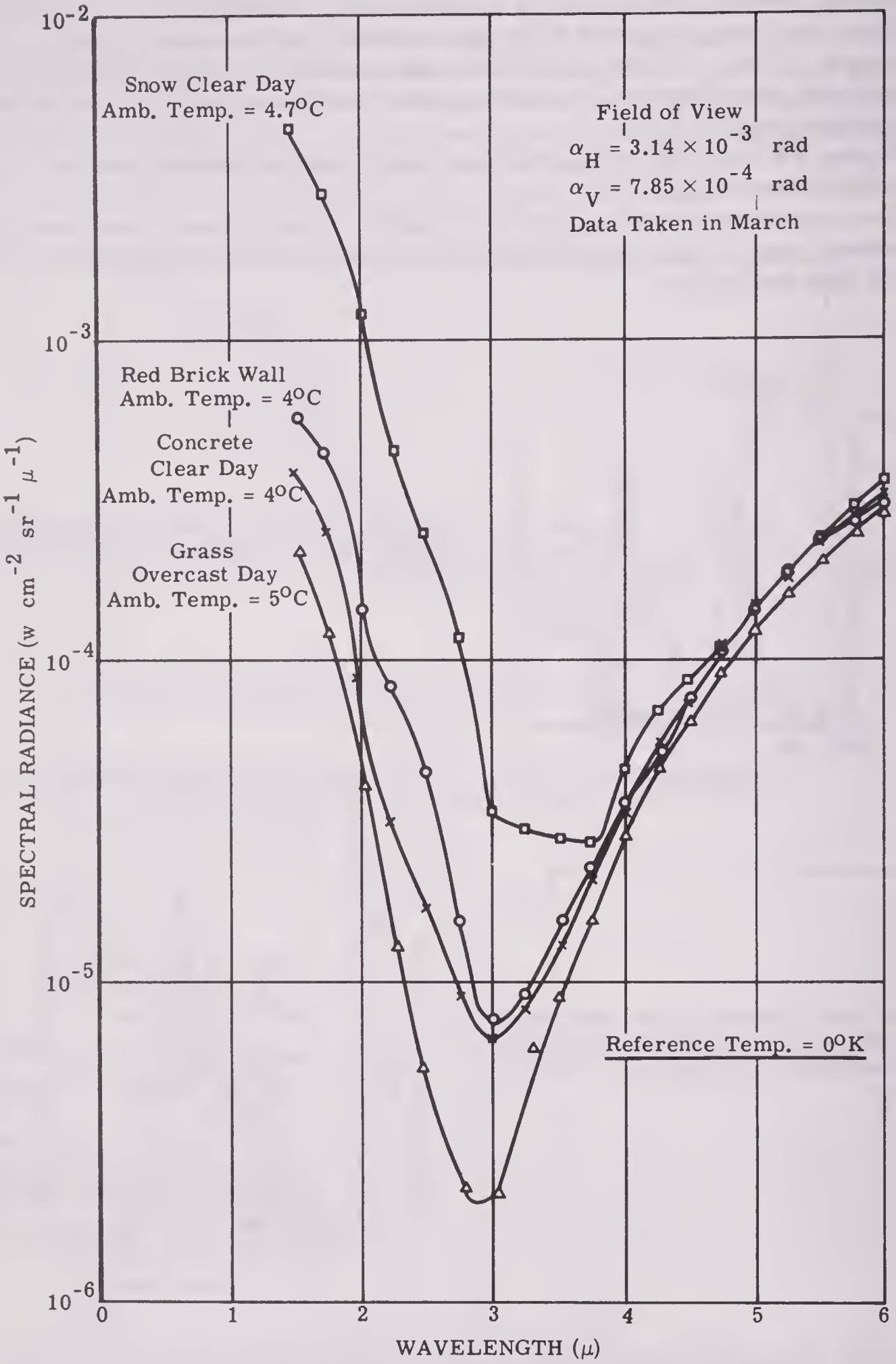


FIG. 5-69. Daytime spectral radiance of miscellaneous targets [31].

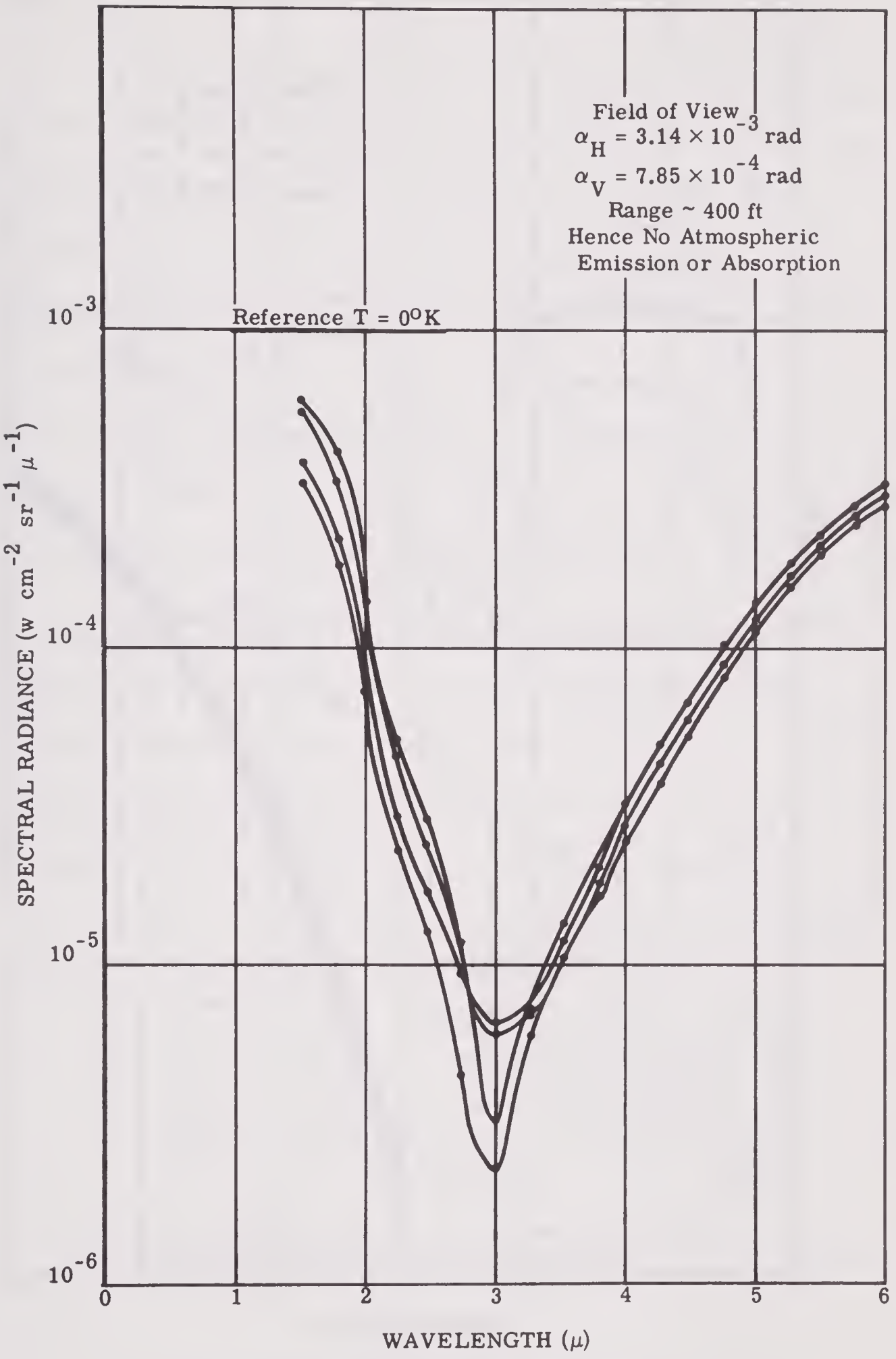


FIG. 5-70. Spectral radiance of concrete, winter day, clear [34].

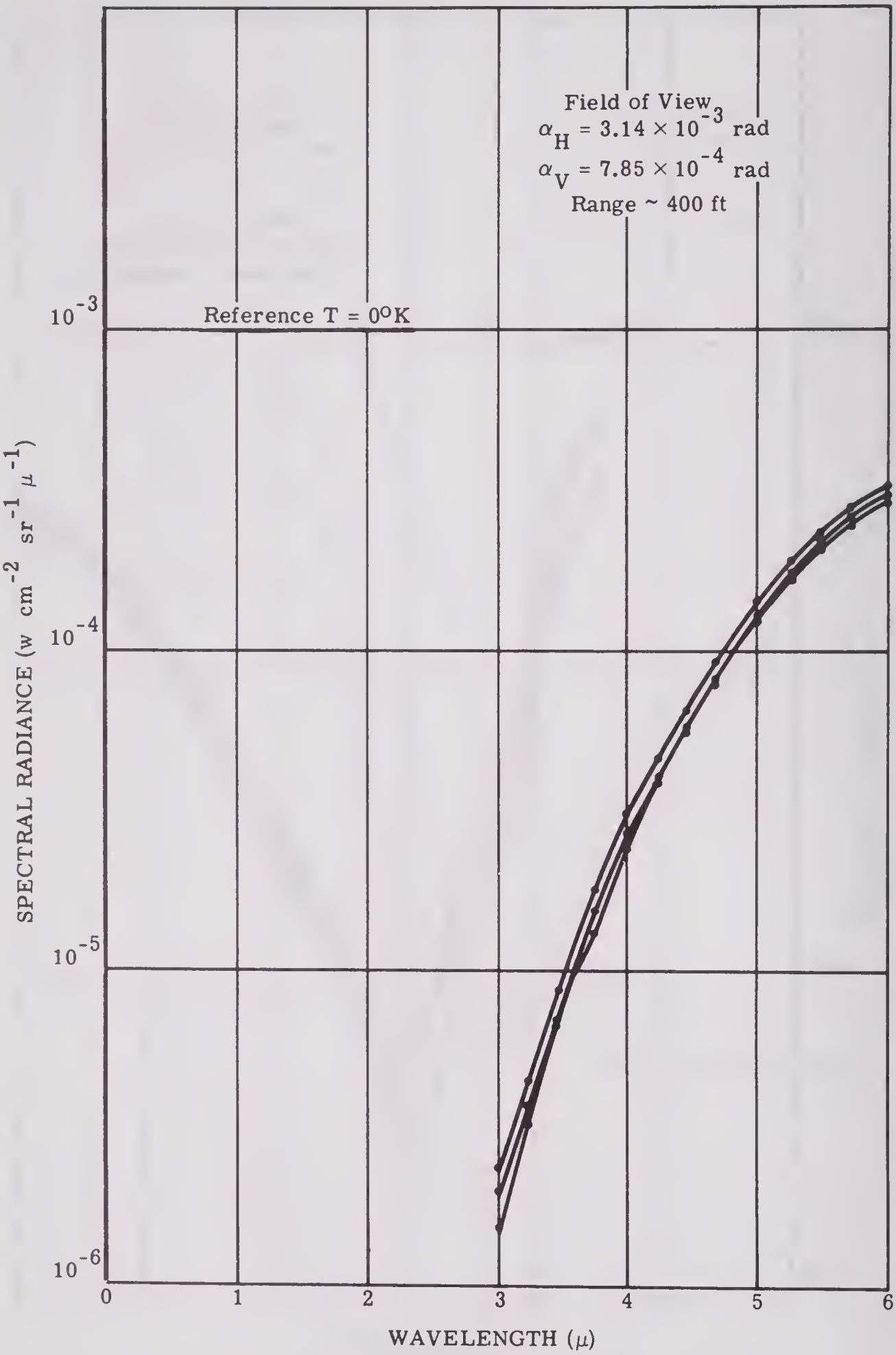
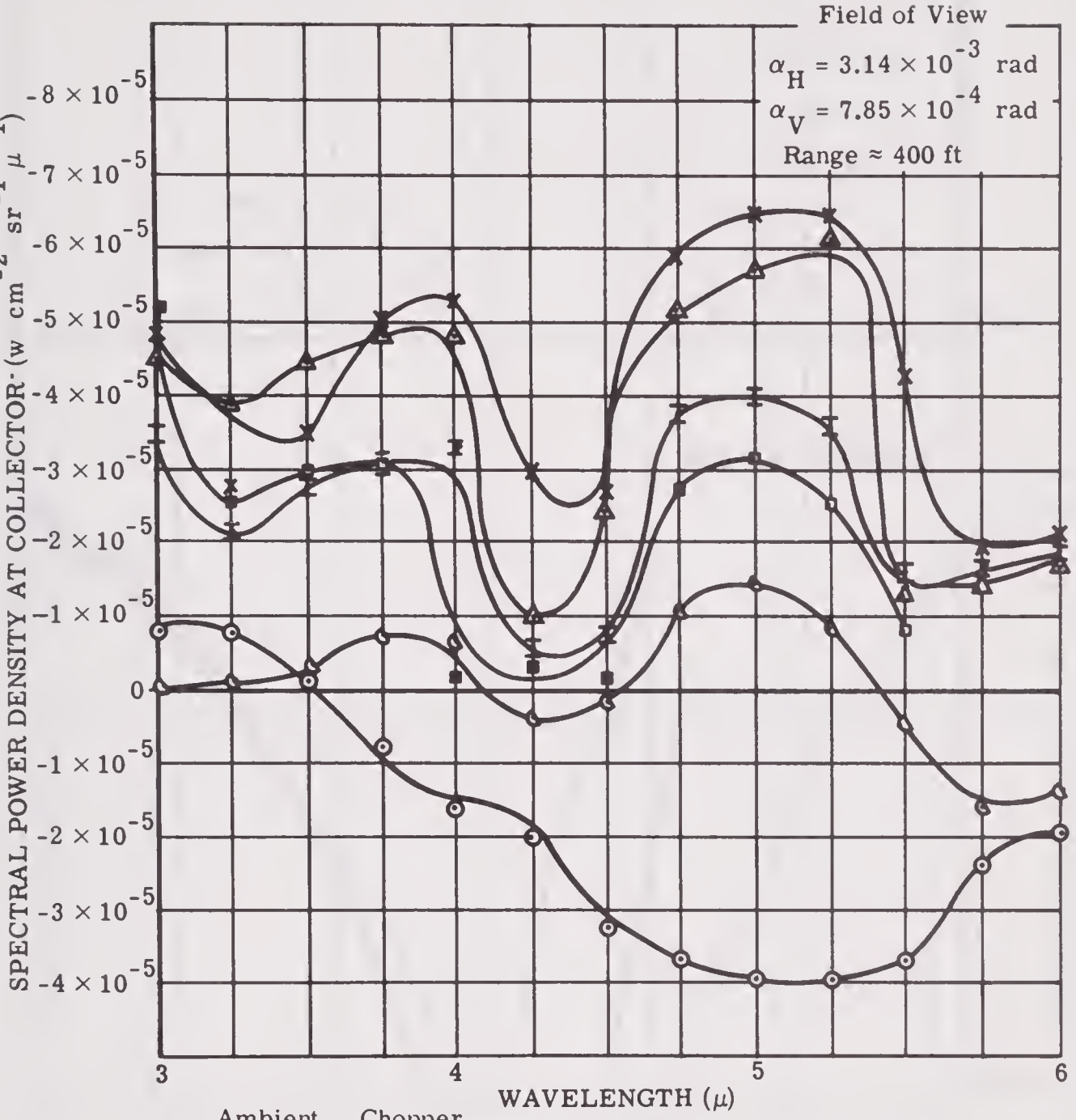


FIG. 5-71. Spectral radiance of concrete, summer night, clear [34].



Run	Ambient Temp.	Chopper Temp.	Illumination	Time	Weather
⊙	25°C	28.8°C	Shade	1:00 P.M.	Sunny, Clear
■	29.2°C	30.8°C	Sun	3:30 P.M.	Sunny, Clear
▲	27.4°C	30.6°C	Sun	5:20 P.M.	Sunny, Clear
×	27°C	30.8°C	Sun	6:25 P.M.	Sunny, Clear
I	27.8°C	30.9°C	Sun	7:15 P.M.	Sunny, Clear
◔	24°C	29.7°C	Shade & Partial Sun	8:05 P.M.	Nearing Sundown Clear

FIG. 5-72. Spectral radiance of concrete wall during an afternoon [34]. Reference temperatures = chopper temperatures.

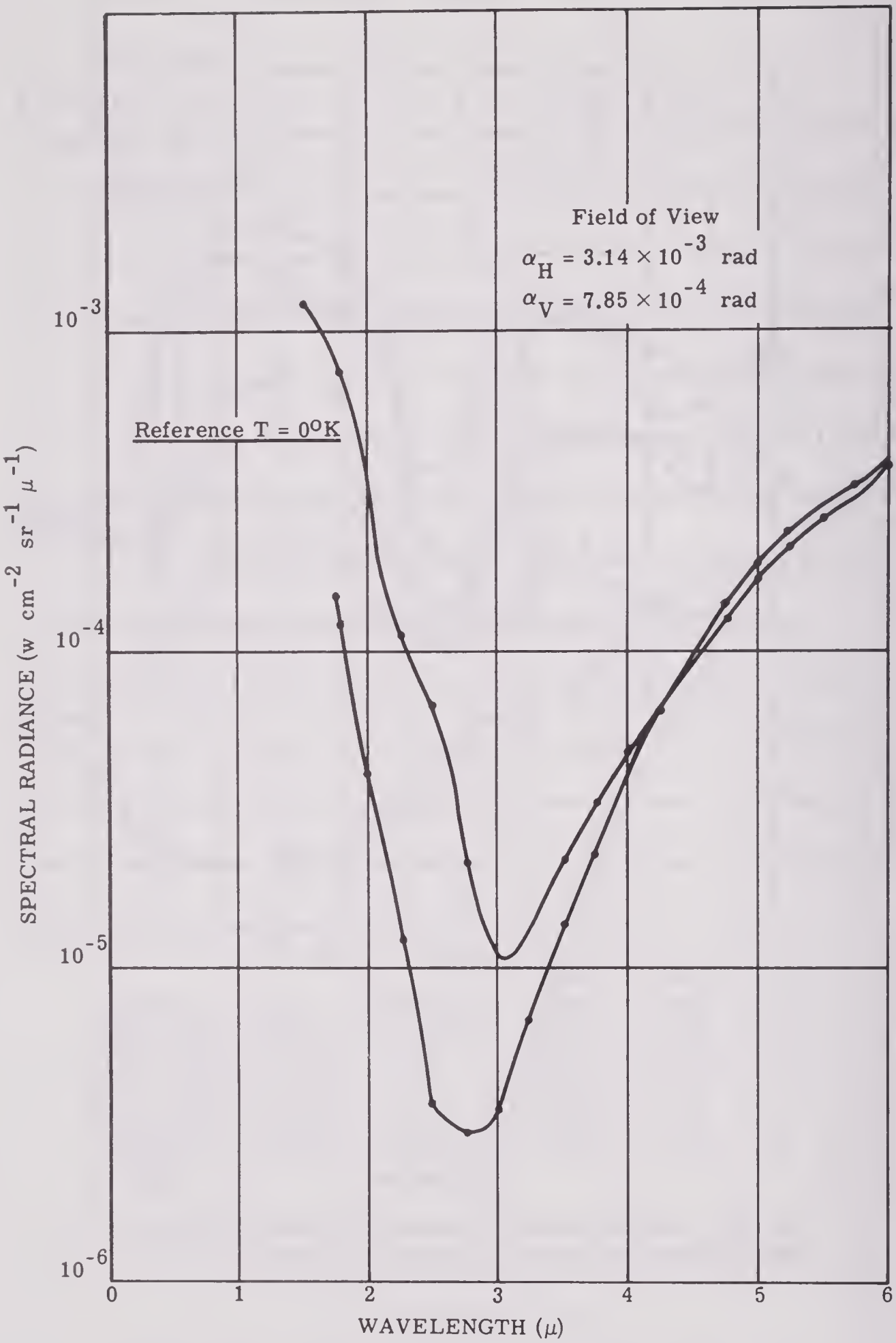


FIG. 5-73. Spectral radiance of concrete, summer day, overcast [34].

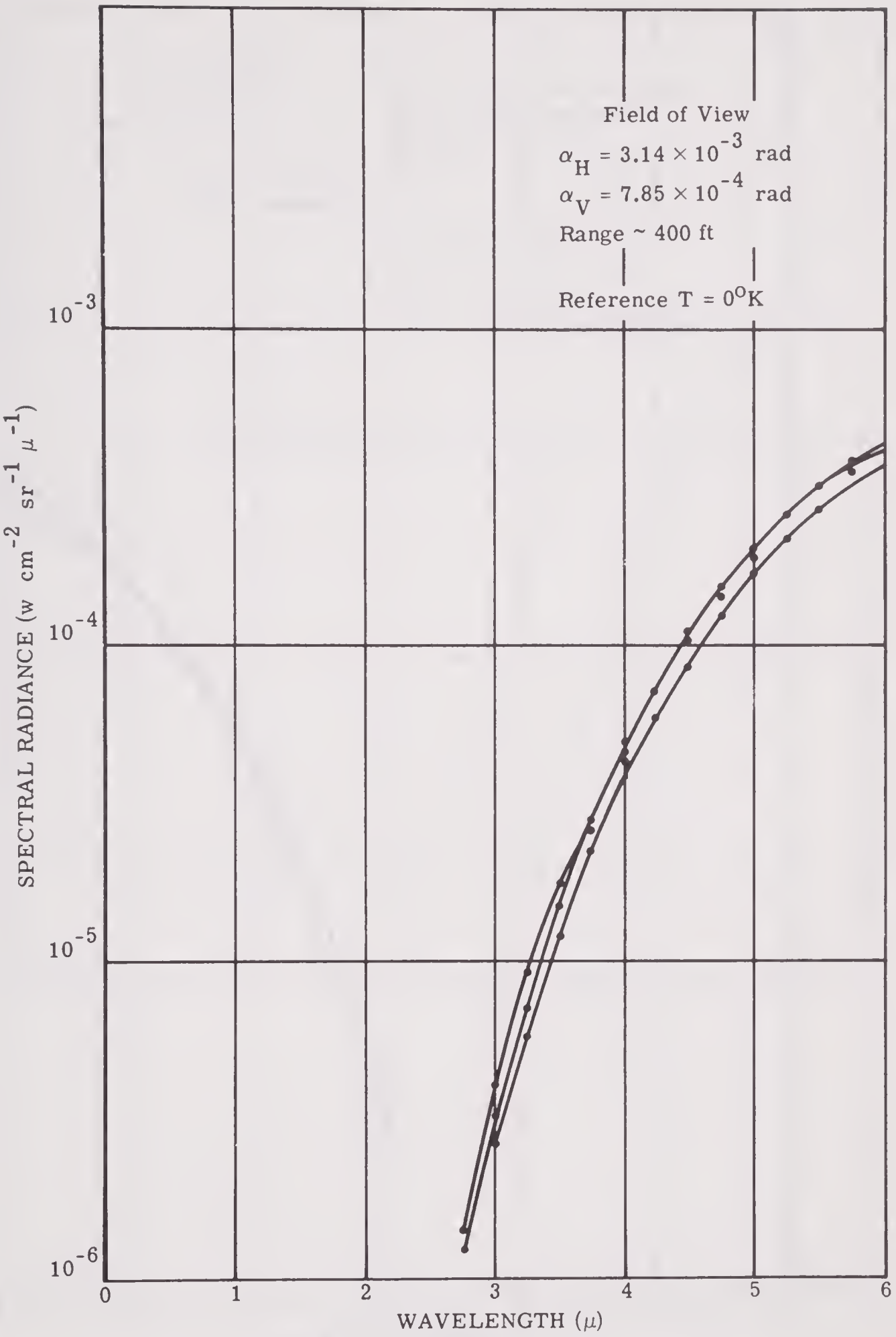


FIG. 5-74. Spectral radiance of concrete, summer night, clear [34].

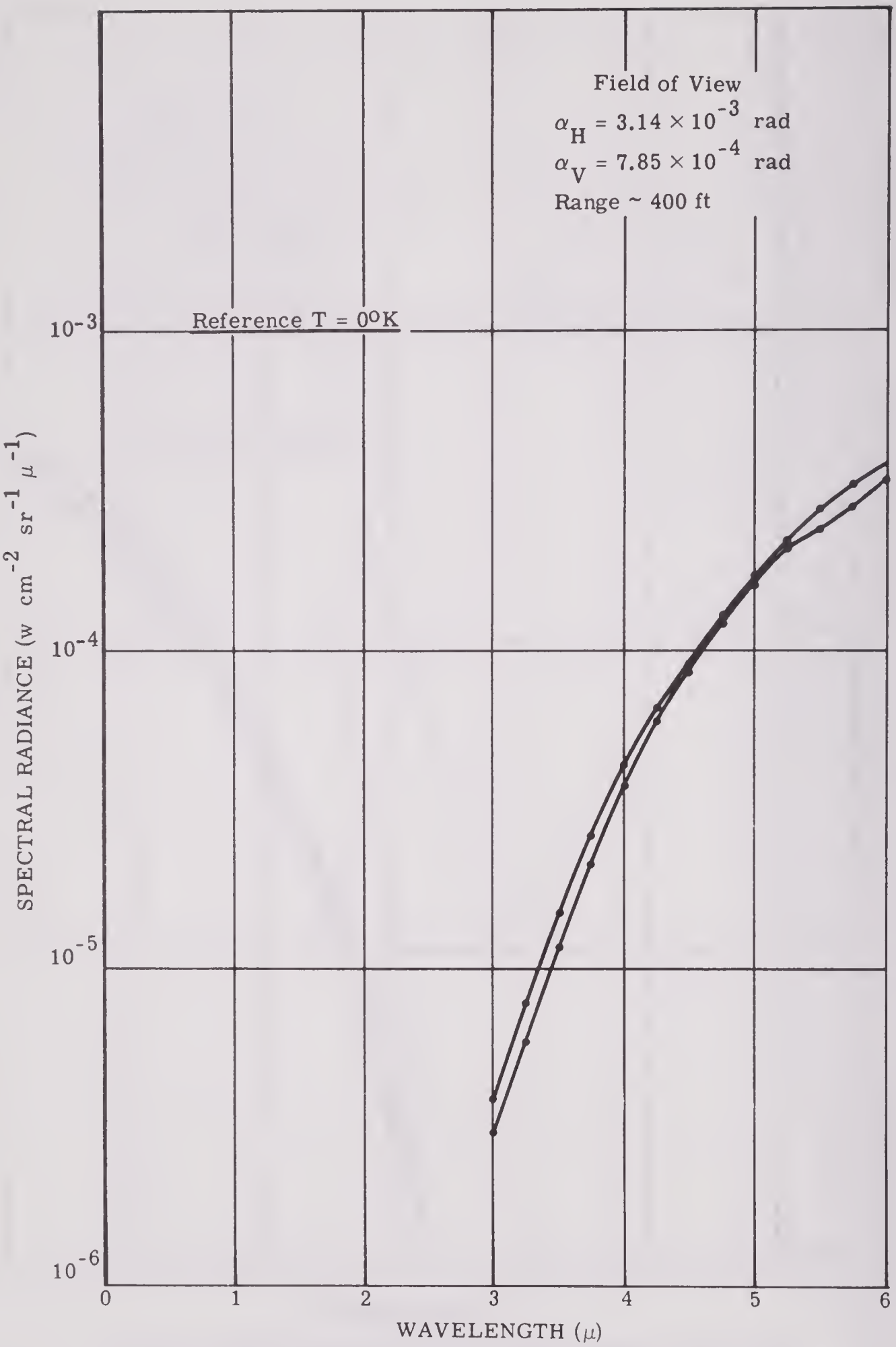


FIG. 5-75. Spectral measurements of comparison of concrete, summer night, overcast [34].

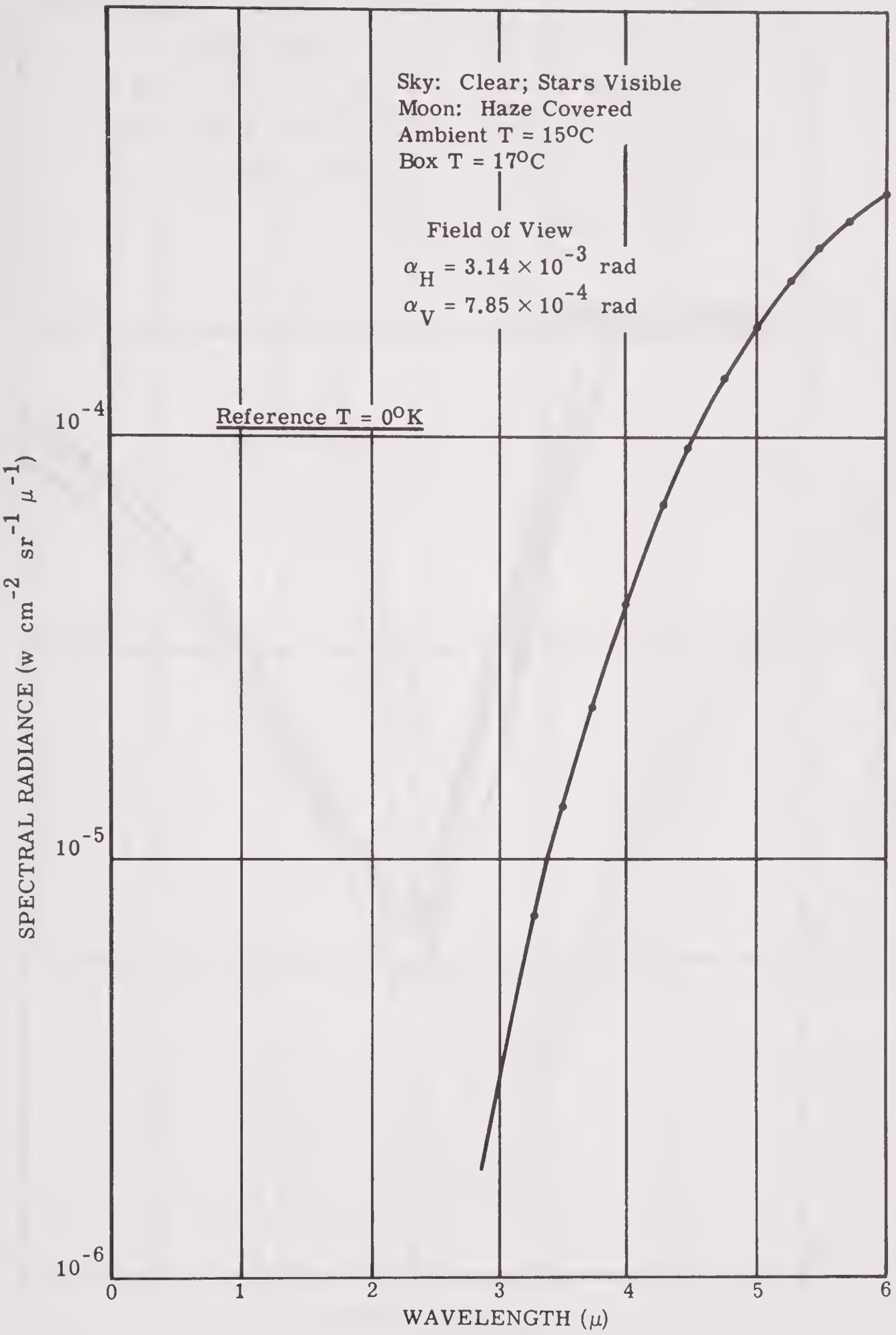


FIG. 5-76. Spectral radiance of concrete at night [34].

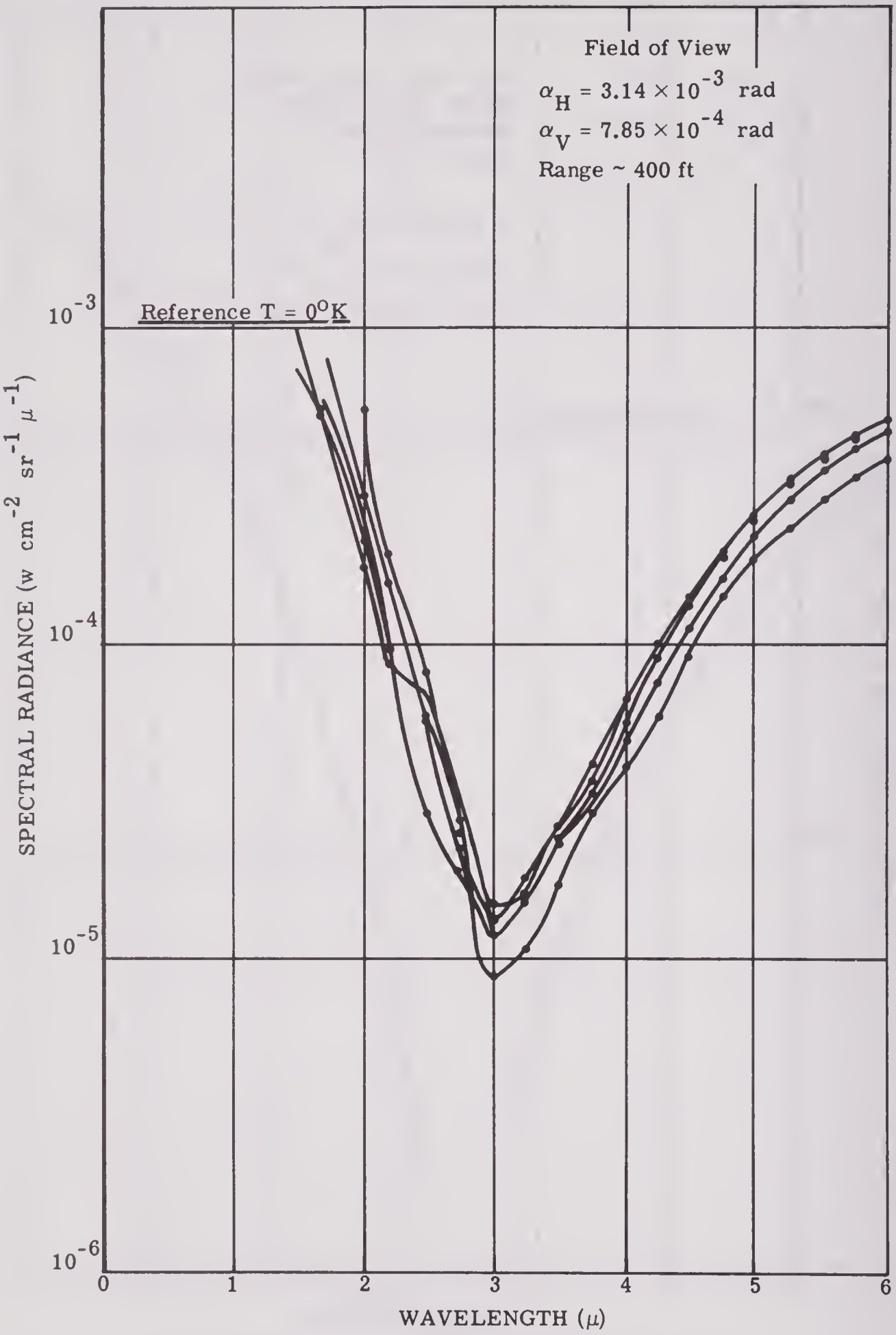


FIG. 5-77. Spectral radiance of damp concrete [34].

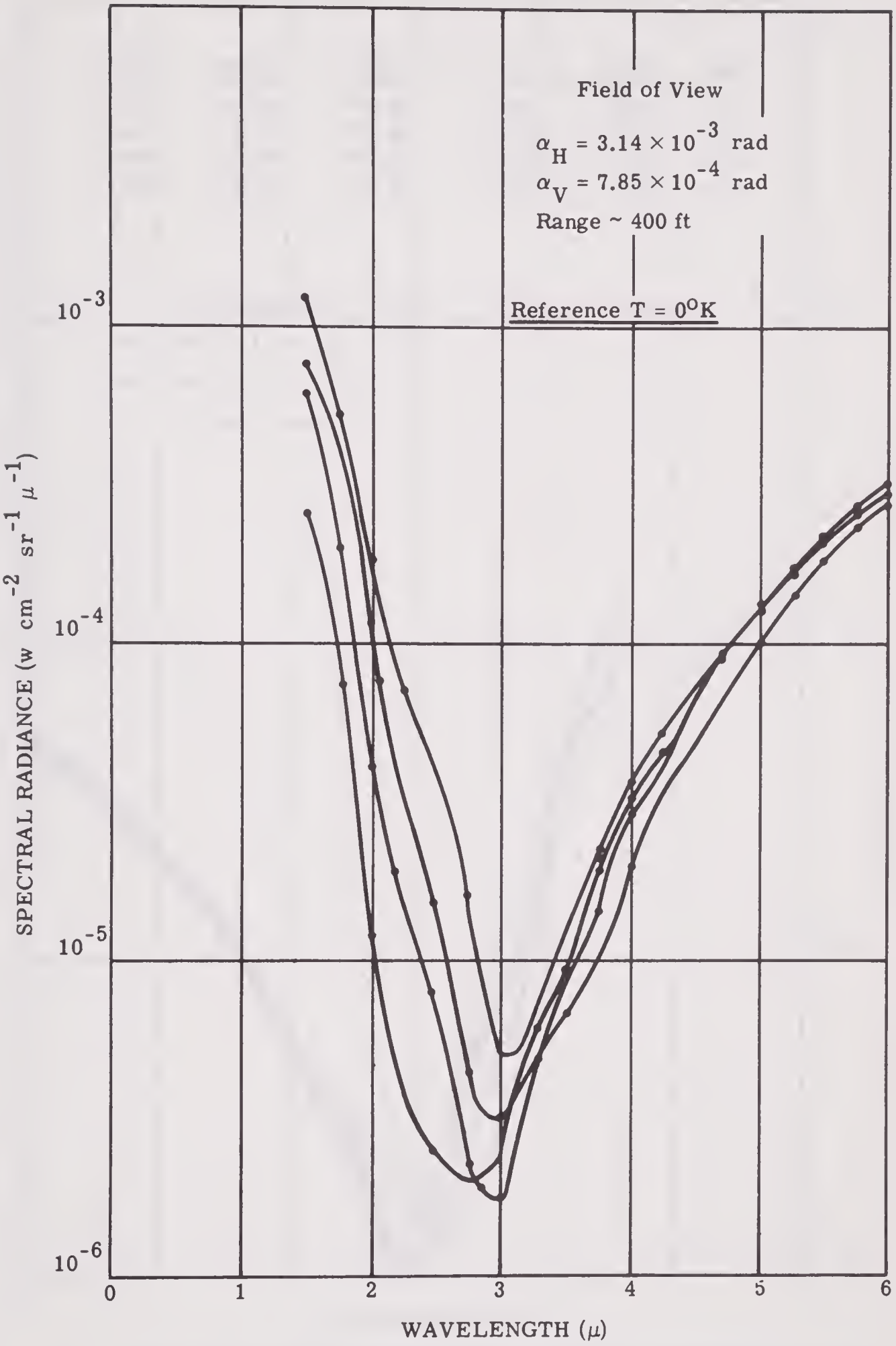


FIG. 5-78. Spectral radiance of concrete through falling snow [34].

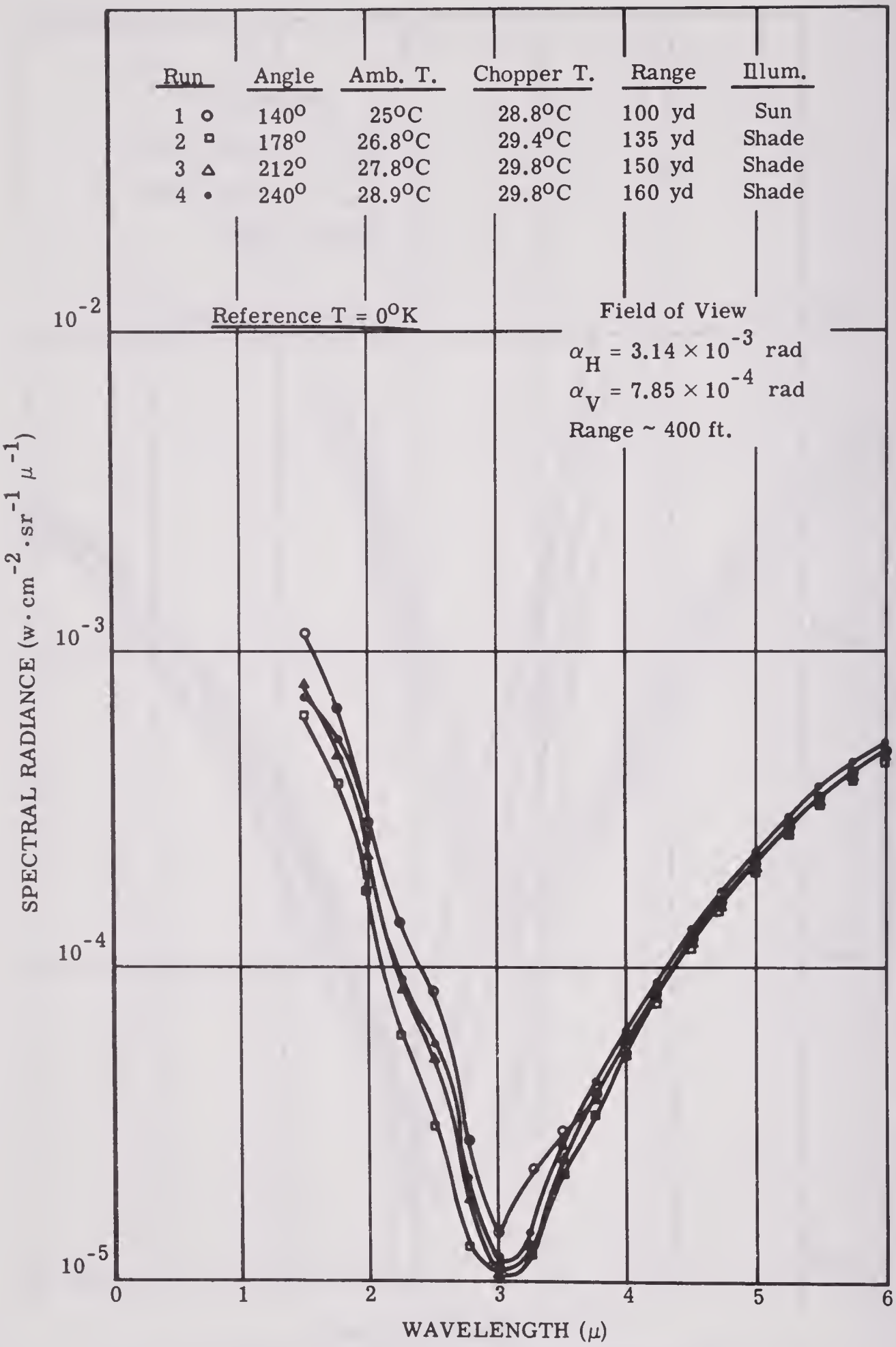


FIG. 5-79. Spectral radiance of concrete wall from four different angles [34].

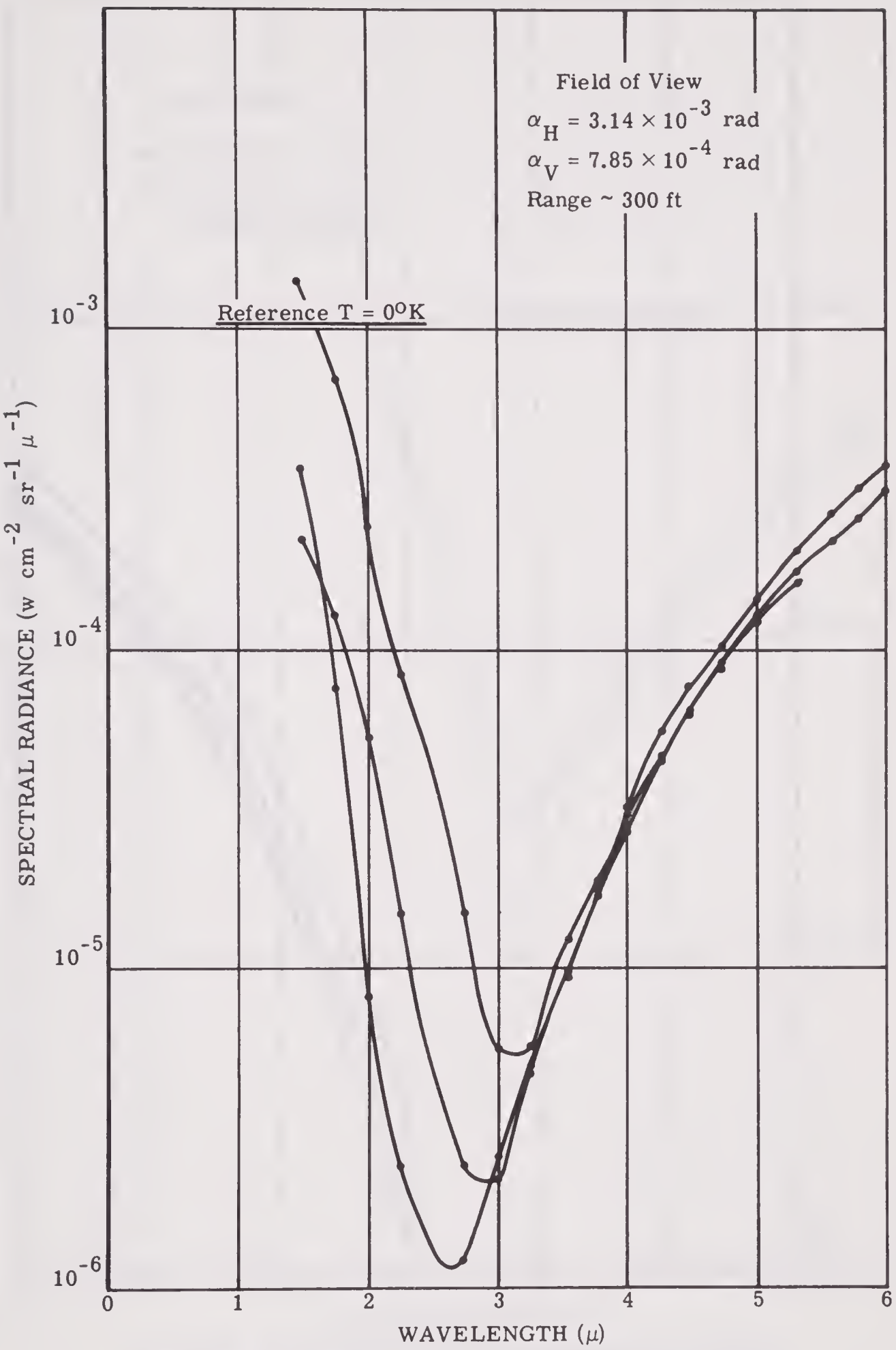


FIG. 5-80. Spectral radiance of grass, winter day, overcast [34].

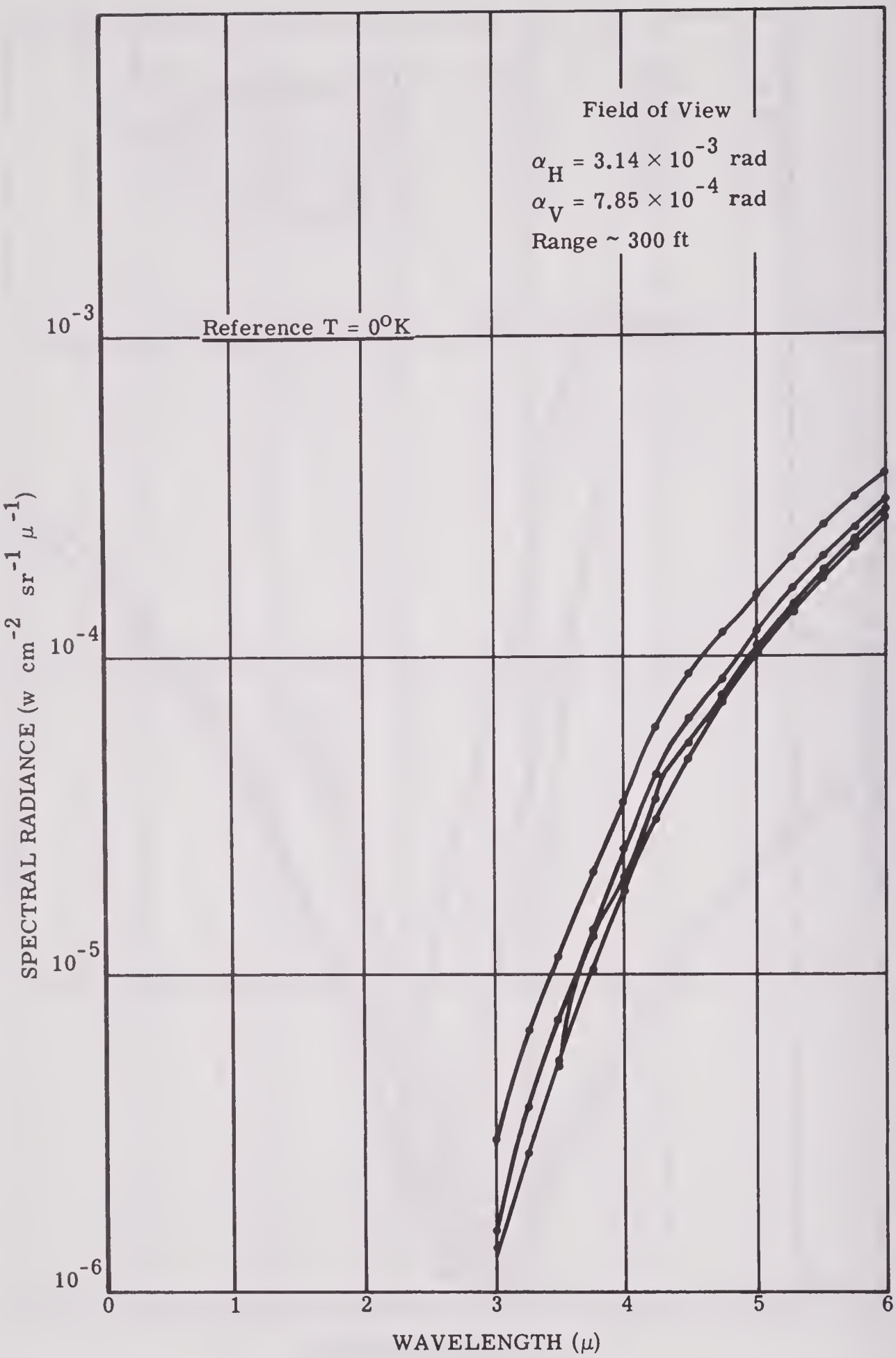


FIG. 5-81. Spectral radiance of grass, winter night, clear [34].

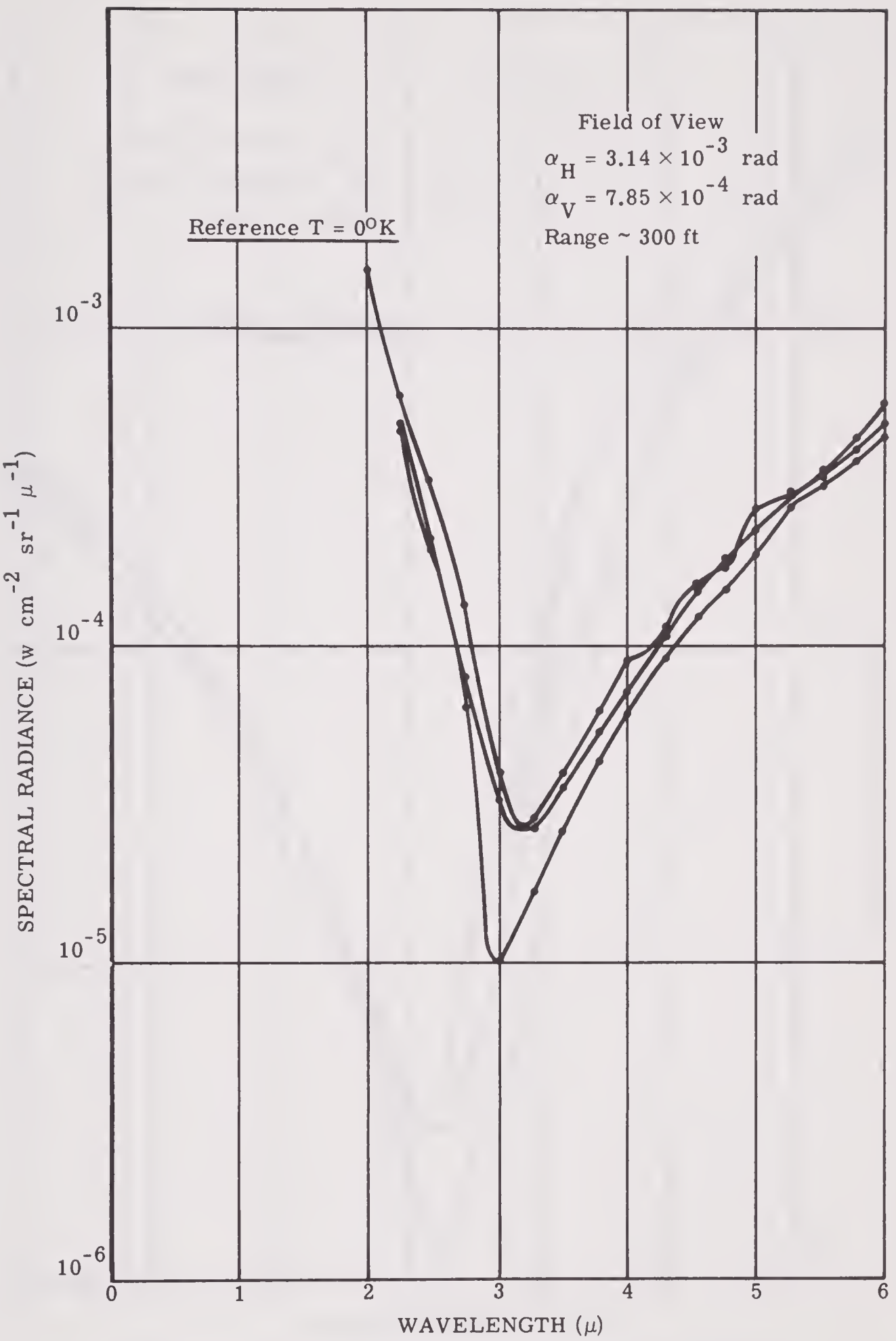


FIG. 5-82. Spectral radiance of grass, summer day, clear [34].

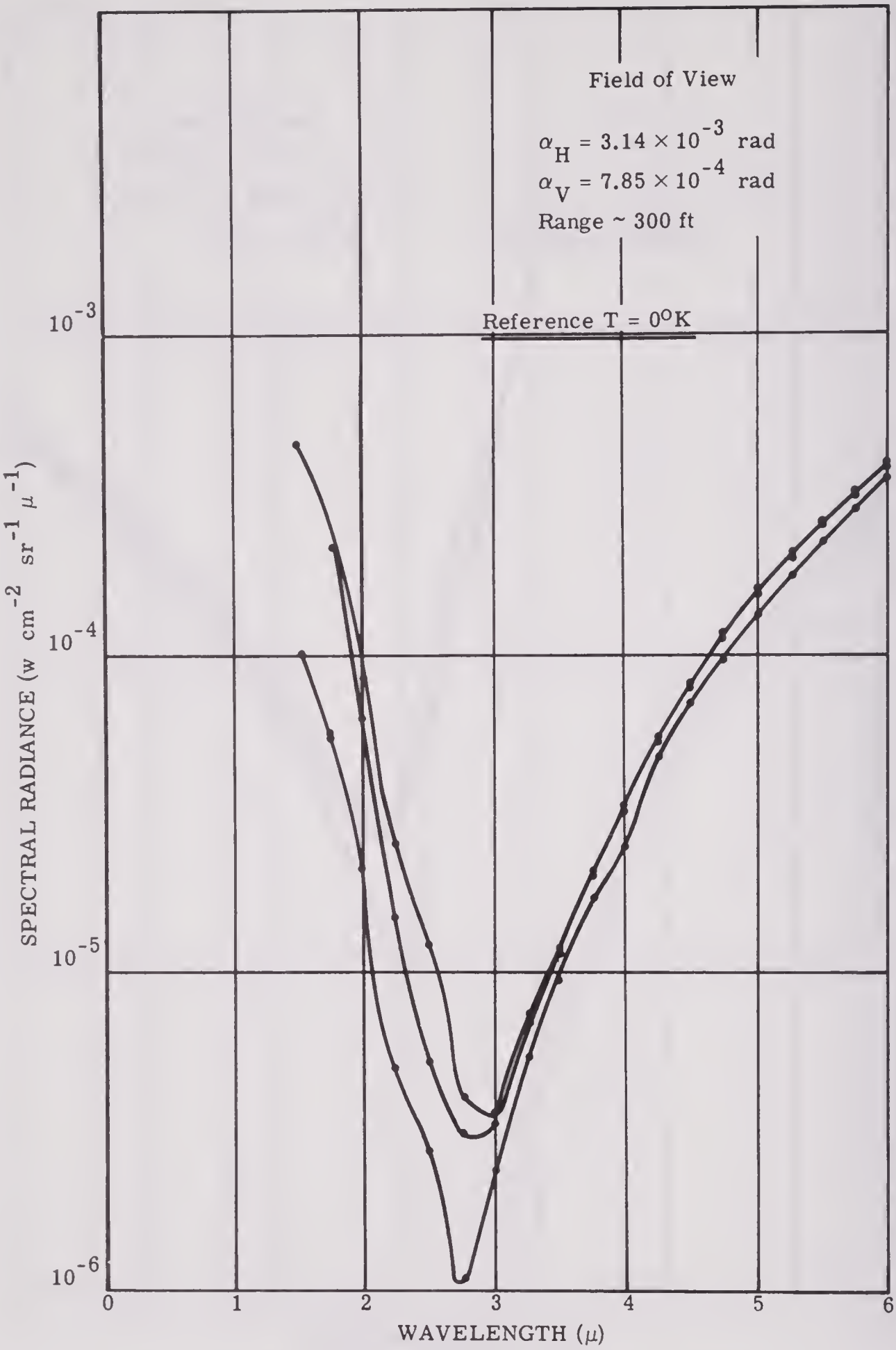


FIG. 5-83. Spectral radiance of grass, summer day, overcast [34].

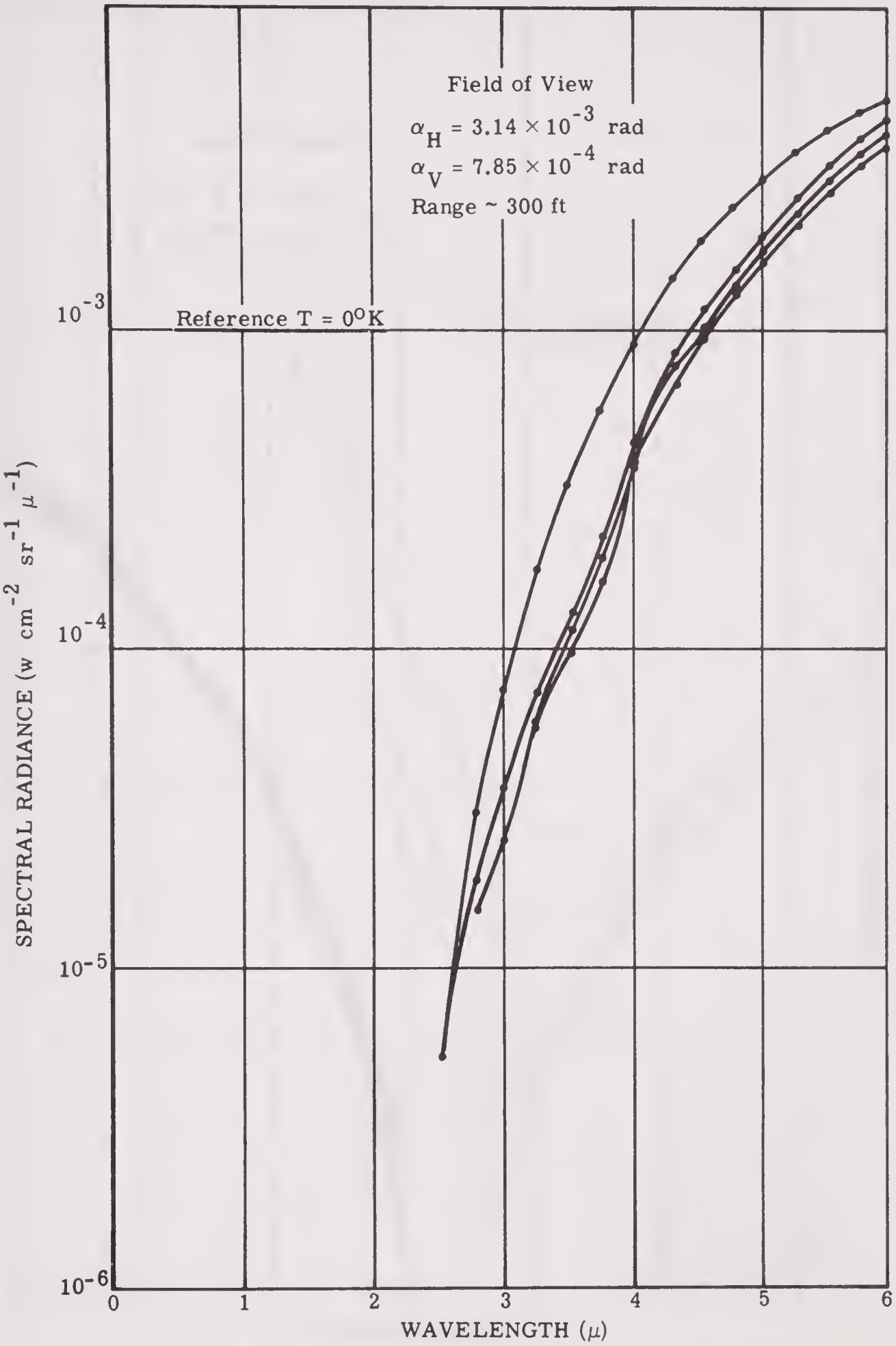


FIG. 5-84. Spectral radiance of grass, summer night, clear [34].

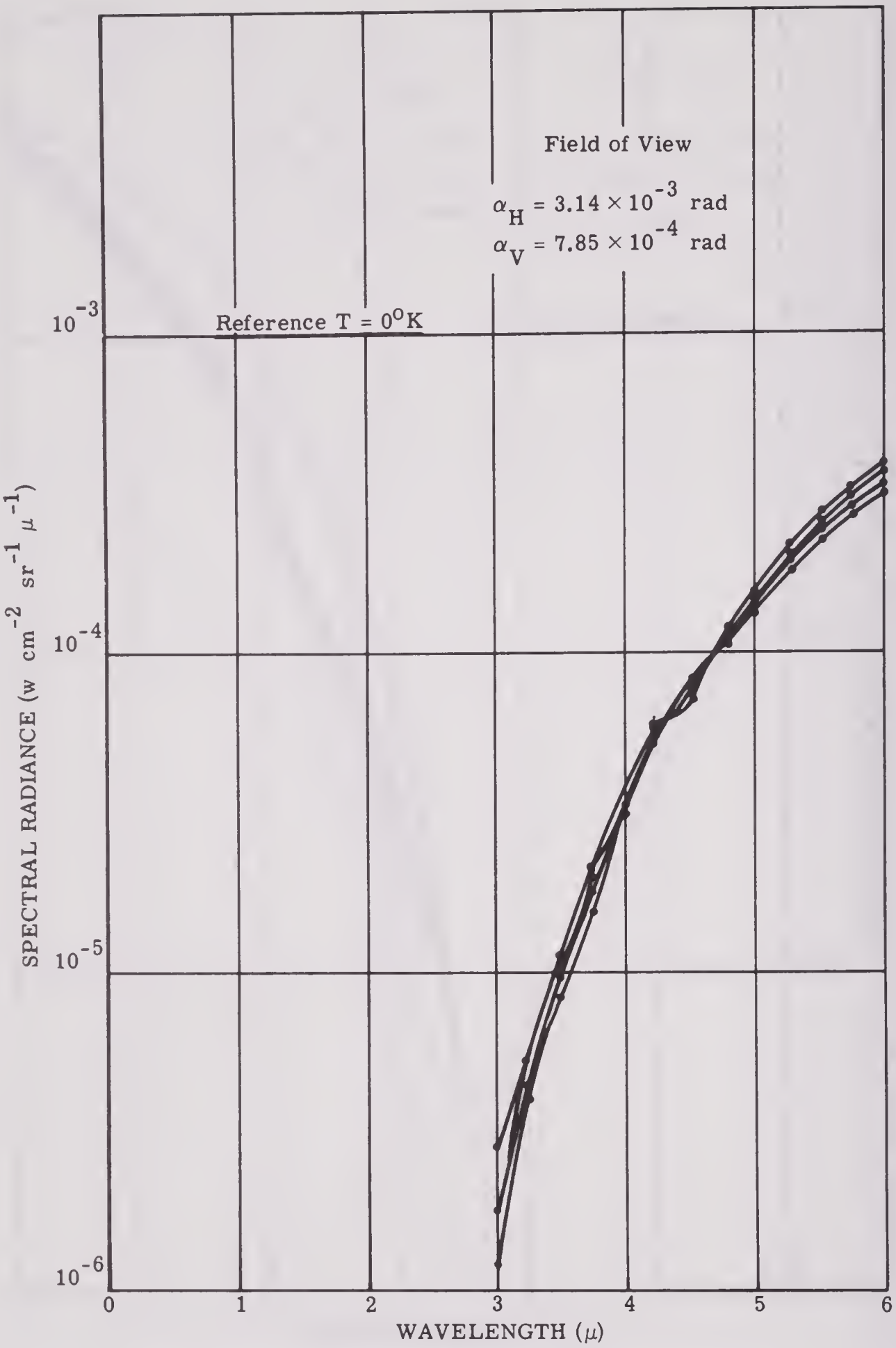


FIG. 5-85. Spectral radiance of grass, summer night, overcast [34].

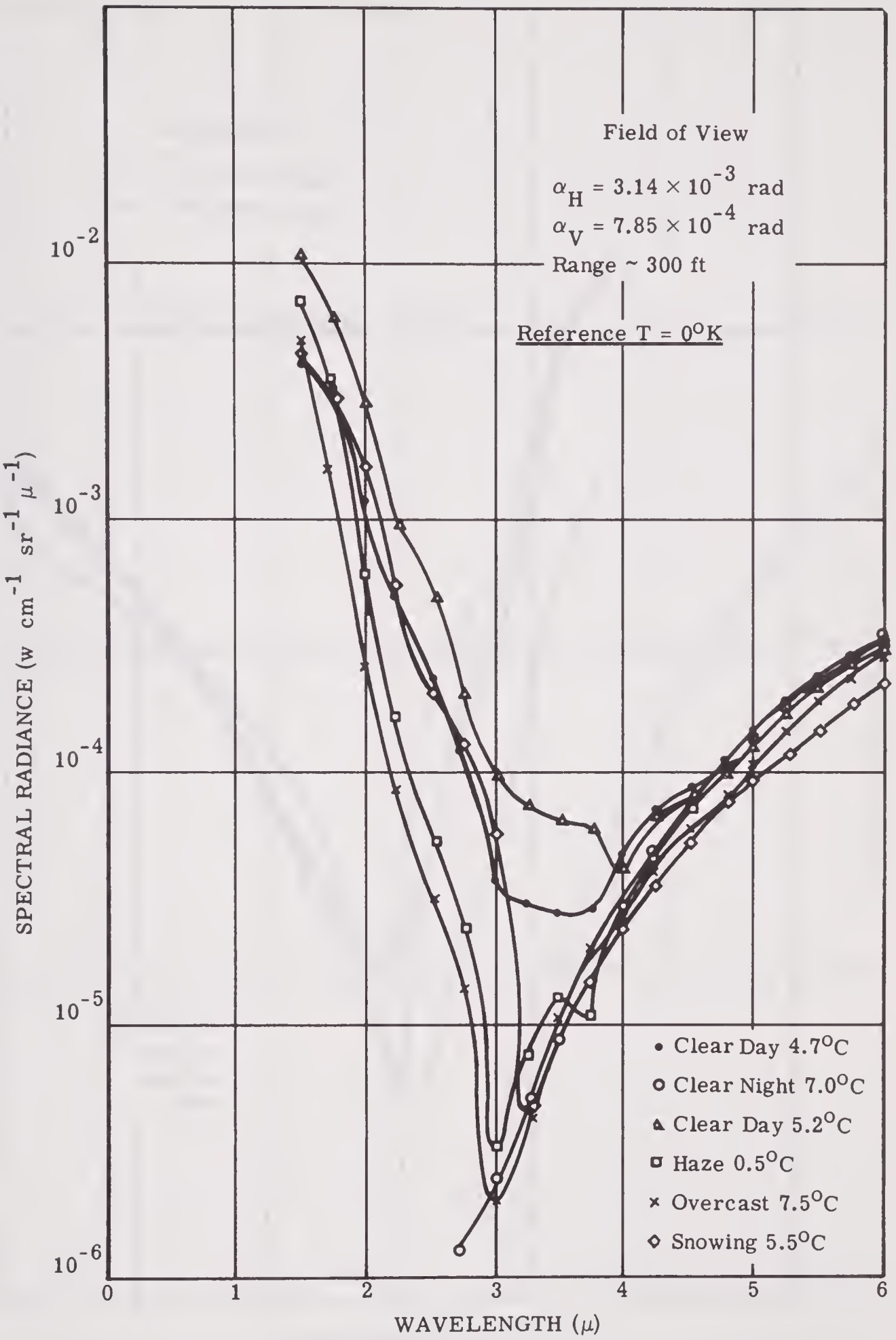


FIG. 5-86. Spectral radiance of snow cover under various weather conditions at various ambient temperatures [34]. Note the high intensity of scattered sunlight.

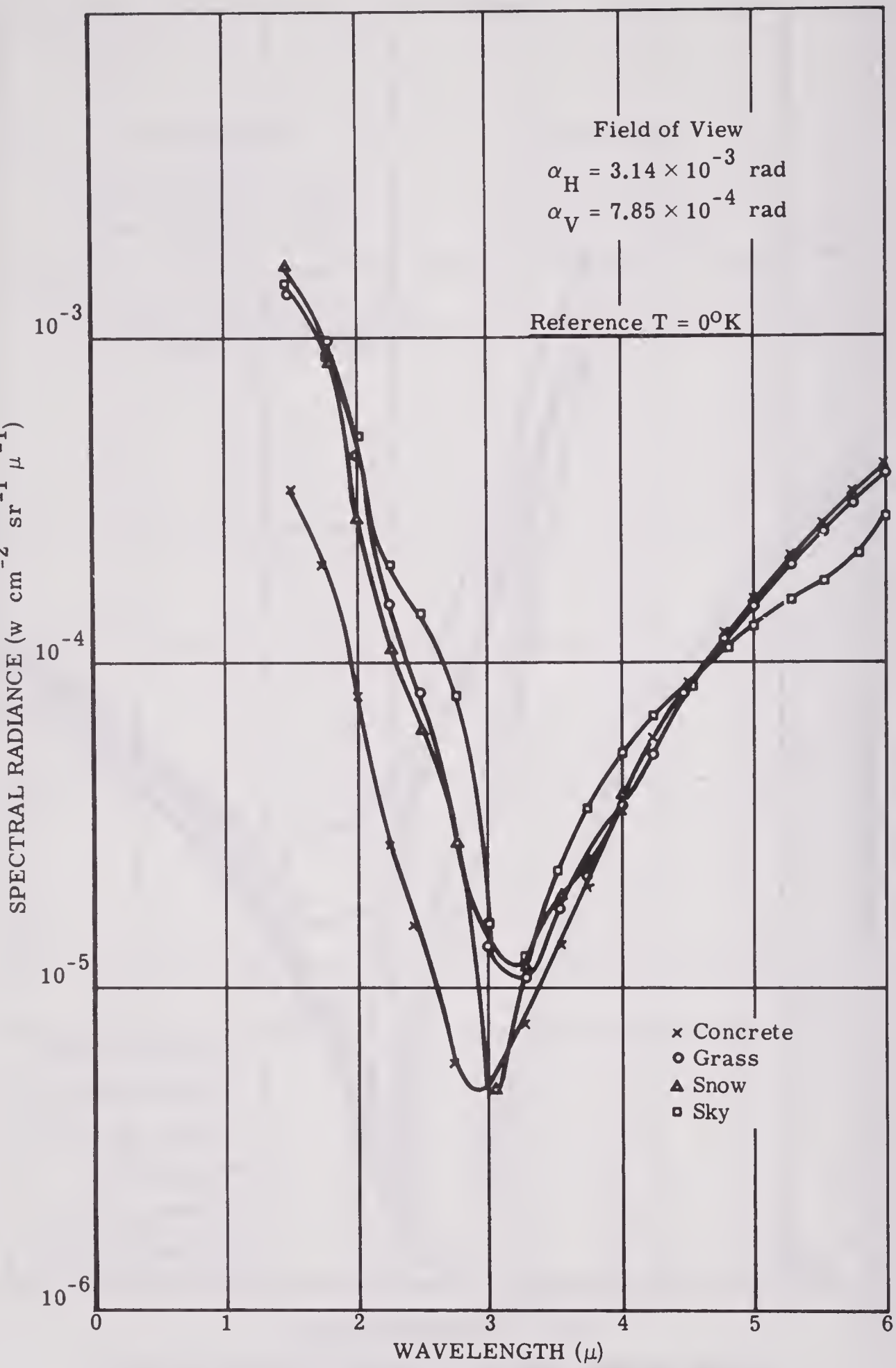


FIG. 5-87. Spectral radiance of sky, concrete, snow, and grass, winter day [34].

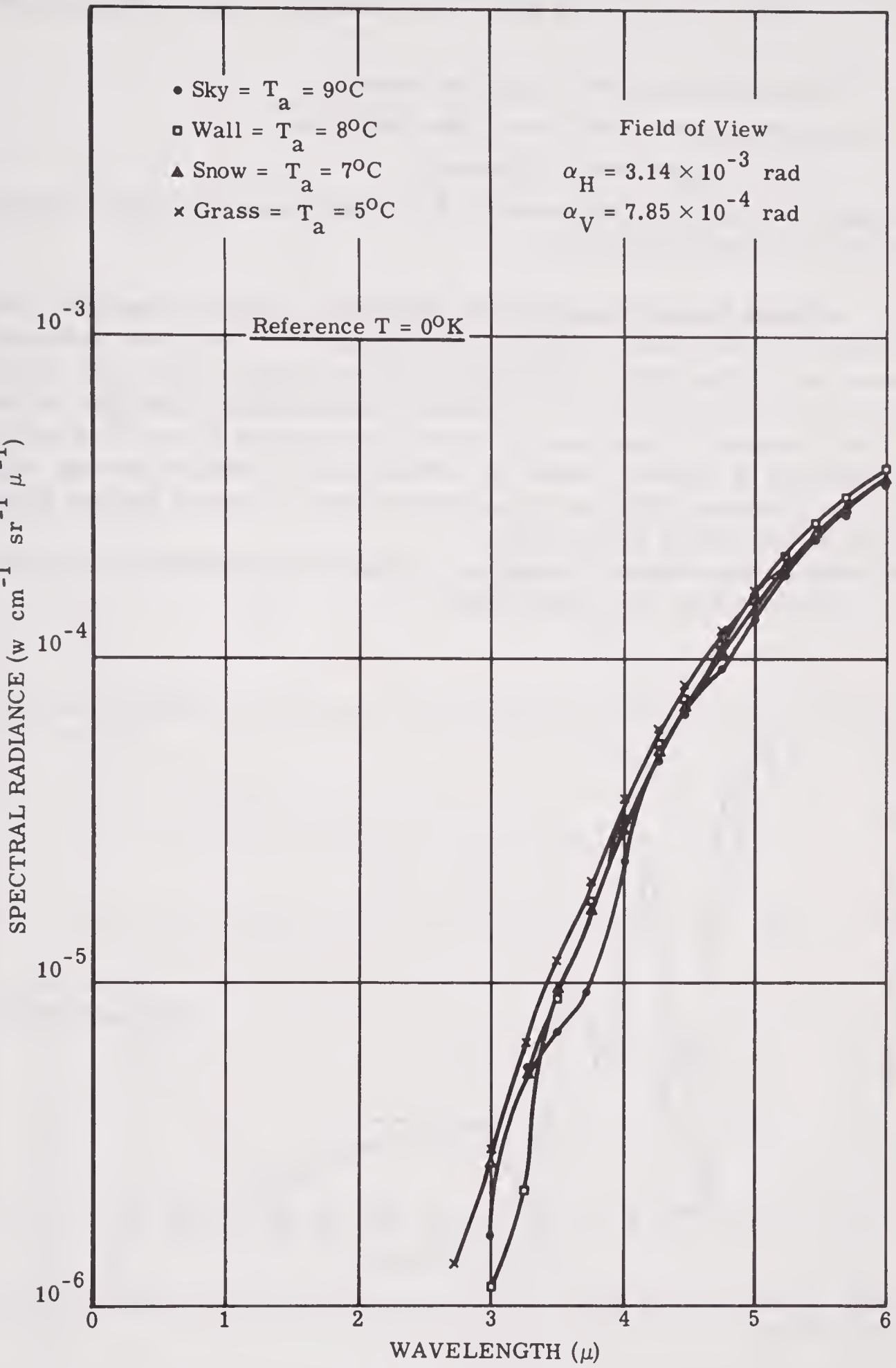


FIG. 5-88. Spectral radiance of sky, concrete, snow, and grass, winter night [34].

5.9. Marine Backgrounds

The radiance of the sea surface at night is the sum of its thermal emission and reflected sky radiance. Factors that determine the character of the marine background are:

1. The infrared optical properties of sea water.
2. Sea-surface geometry and wave-slope distribution.
3. Sea-surface temperature distribution.

Atmospheric transmission and emission in the optical path from scene to observing instrument is covered in Chapter 6.

5.9.1. Infrared Optical Properties of Sea Water. Water is essentially opaque to infrared radiation longer than 3 μ. Few liquids have absorption coefficients of the same order of magnitude. Consequently, the sea surface, which is 0.01 cm thick, determines the radiance of the sea. Subsurface scattering of sky radiation is absent. The optical influence of thin layers of surface contamination is negligible except for the suppression of capillary waves by surface tension changes—causing "slicks." There is no significant difference in the transmissivity of sea and distilled water for these thin layers in the 2- to 15-μ region.

The infrared transmissivity, reflectivity, emissivity and indices of refraction for water are shown in Figs. 5-89 to 5-92 [36,37].

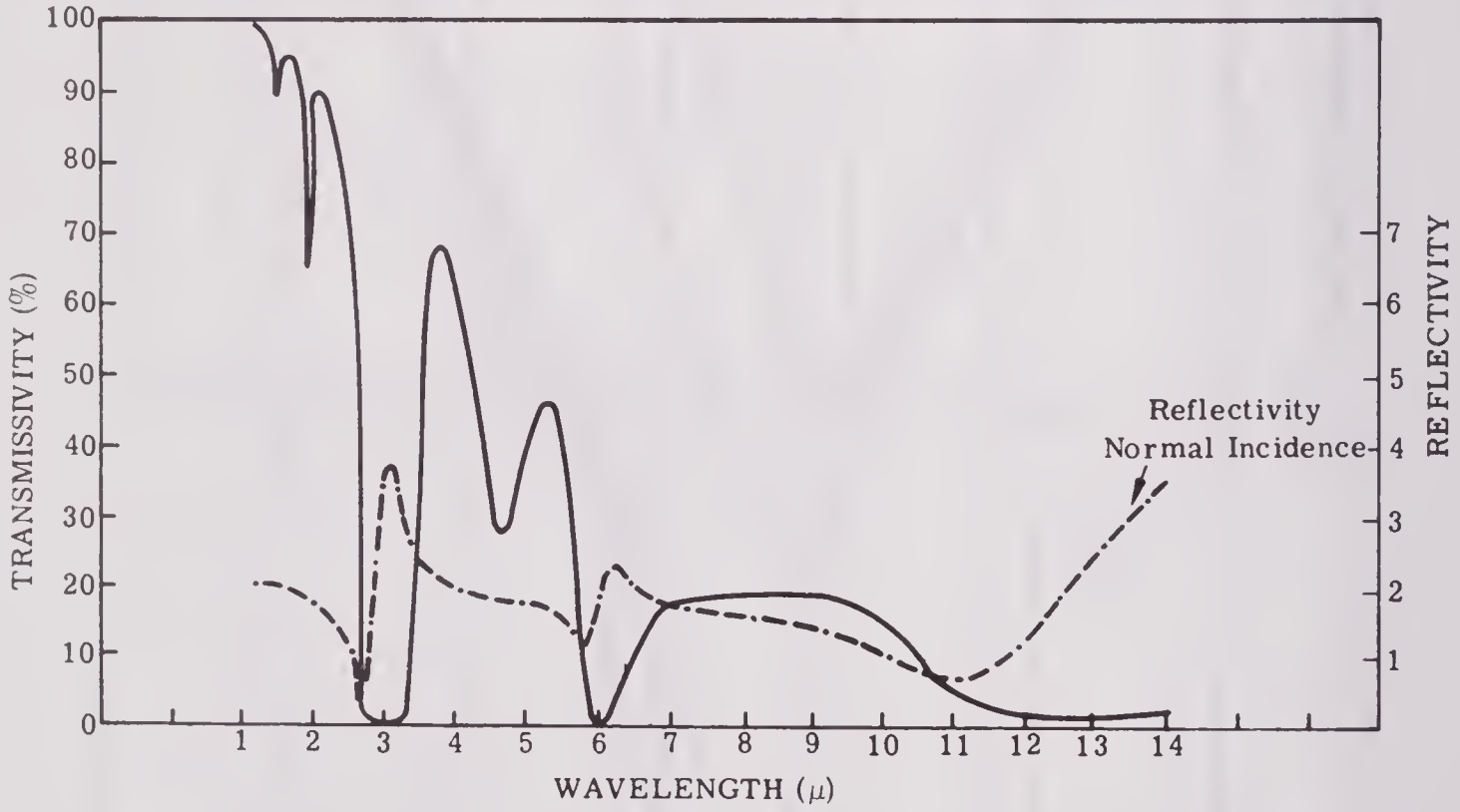


FIG. 5-89. Transmissivity of 0.003 cm of sea water and reflectivity of a free sea-water surface (dashed line) [36].

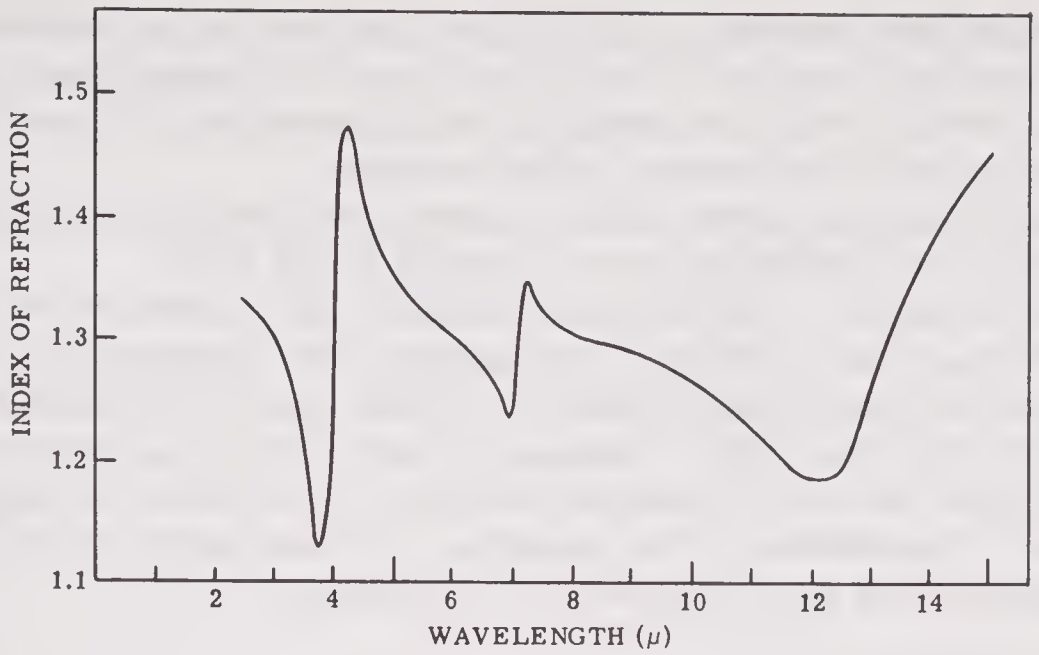


FIG. 5-90. Indices of refraction of water calculated from reflectivity data in Fig. 5-89.

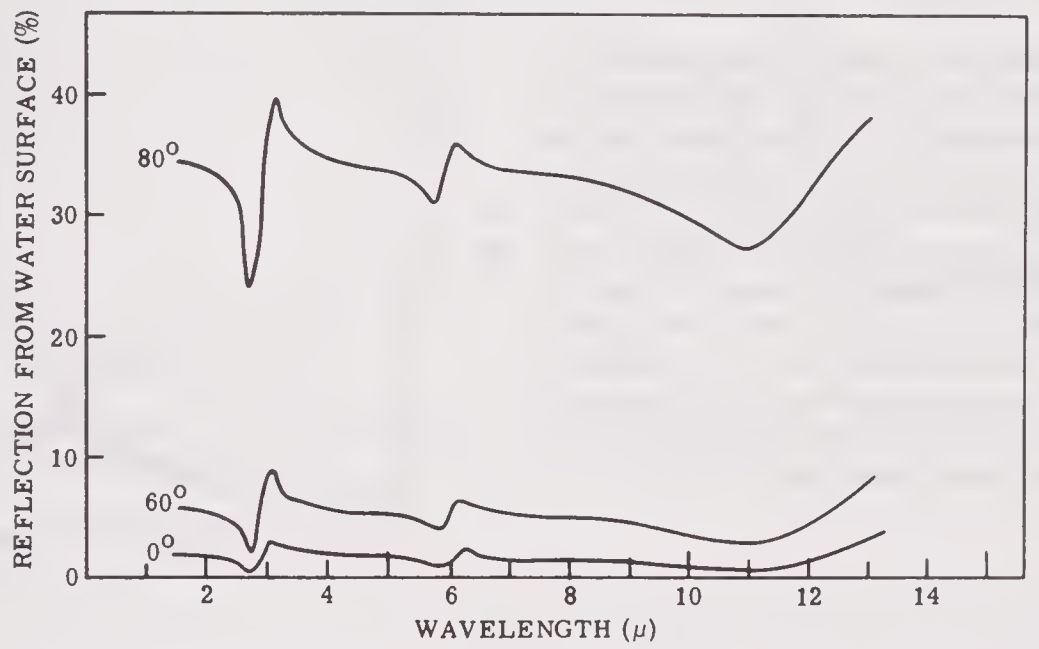


FIG. 5-91. Reflection from a water surface at 0°, 60°, and 80° angle of incidence calculated from data in Fig. 5-90.

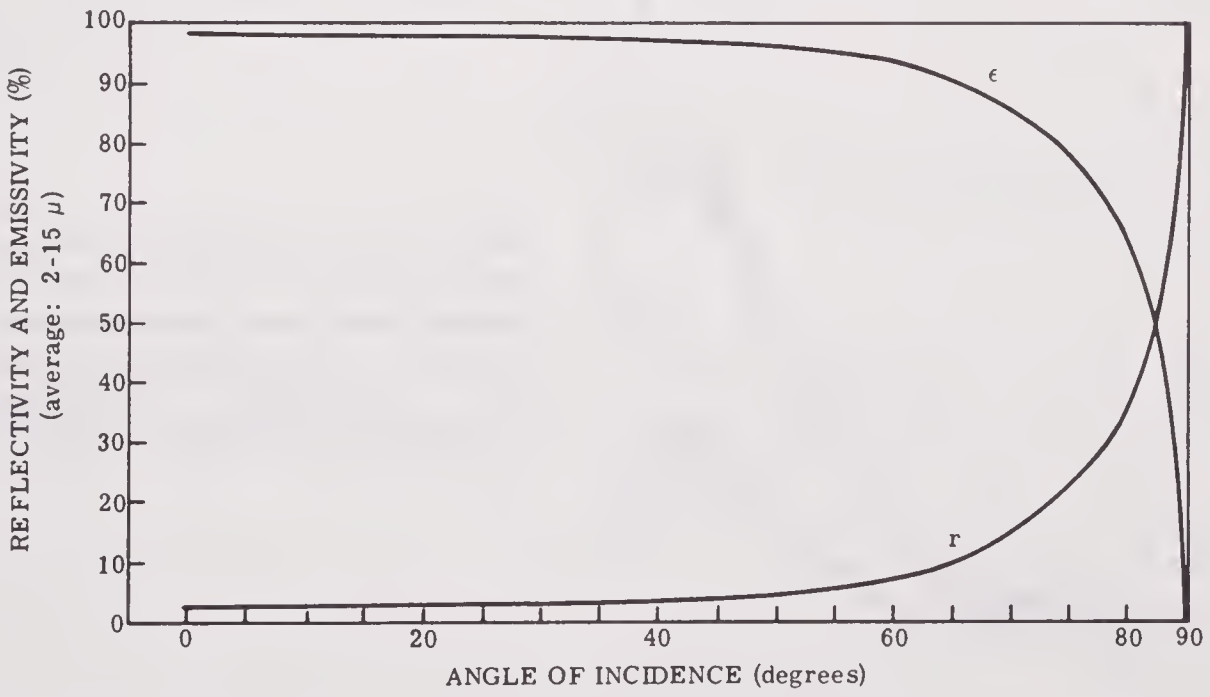


FIG. 5-92. Reflectivity and emissivity of water (2- to 15-μ average) *versus* angle of incidence, calculated from averaged data of Fig. 5-90; see [37]. Note scale change.

5.9.2. Sea-Surface Geometry. The effect of wave slope on the reflectivity of a sea surface roughened by a Beaufort 4 wind (11 to 16 knots, white caps) is seen in Fig. 5-93. Here, for an average rough sea, the reflectivity approaches 20% near the horizon. Consequently, the emissivity remains at 80% or higher.

The radiance of the sea surface along an azimuth 90° from that of the sun, in daylight for clear and for overcast conditions, is shown in Fig. 5-94.

Information is lacking on similar observations for the radiance of the sea surface at night. However, the variation of sky radiance with zenith angle is similar day and night, and the photographic reflectivity is about equal to the average for the infrared from 2 to 15 μ (Fig. 5-92). Consequently, the curves in Fig. 5-94 are instructive because they show the general shape of that part of the radiance of the sea surface at night due to the reflection of sky radiation. To these curves must be added the infrared radiance of the sea surface due to its temperature.

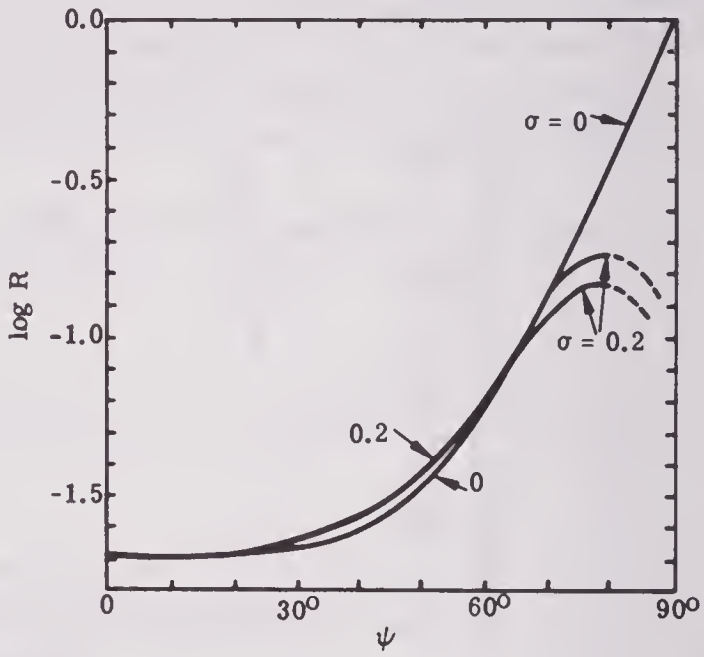


FIG. 5-93. Reflection of solar radiation from a flat surface ($\sigma = 0$) and a surface roughened by a Beaufort 4 wind ($\sigma = 0.2$). The albedo R varies from 0.02 for a zenith sun, $\psi(= 0°)$ to unity for the sun at the horizon ($\psi = 90°$) on a flat sea surface. For a rough surface, shadowing and multiple reflections become important factors when the sun is low. The lower and upper branches of the curve marked $\sigma = 0.2$ represent two assumptions regarding the effect of multiple reflection. True values are expected to lie between the indicated limits [38].

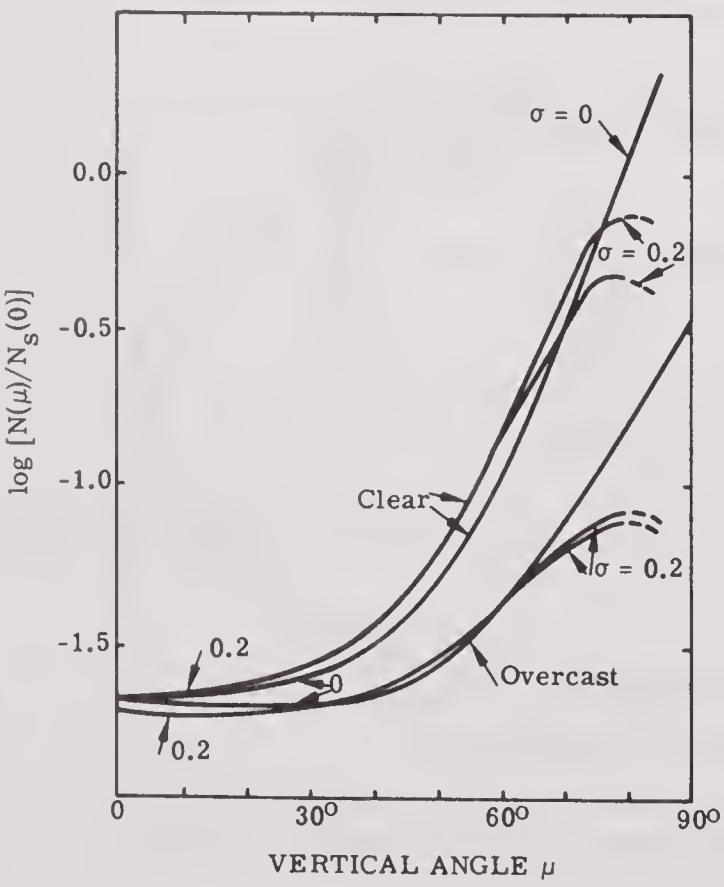


FIG. 5-94. The radiance of the sea surface, $N(\mu)$, divided by the sky radiance at the zenith, $N_s(0)$, as a function of the vertical angle μ. The curves are computed for a flat ($\sigma = 0$) and rough ($\sigma = 0.2$) surface for two of the sky conditions illustrated in Fig. 5-93 [38].

Examples of the spectral radiance of the sea for day and after sundown are shown in Figs. 5-95, 5-96, and 5-97 [1].

For further data on sea-surface geometry see [40].

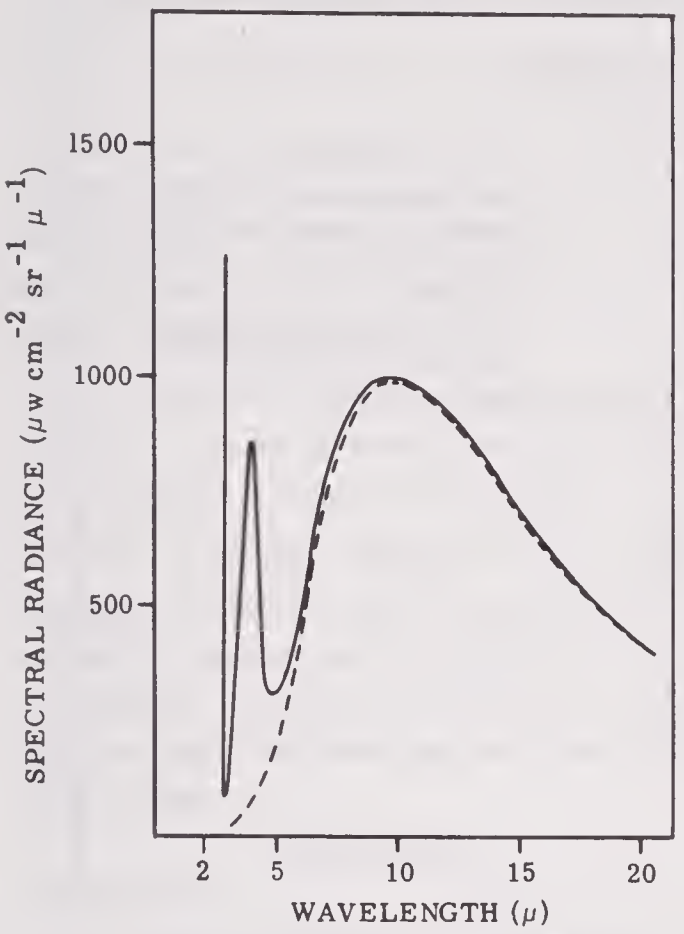


FIG. 5-95. Spectral radiance of the Banana River at Cocoa Beach, Florida [33].

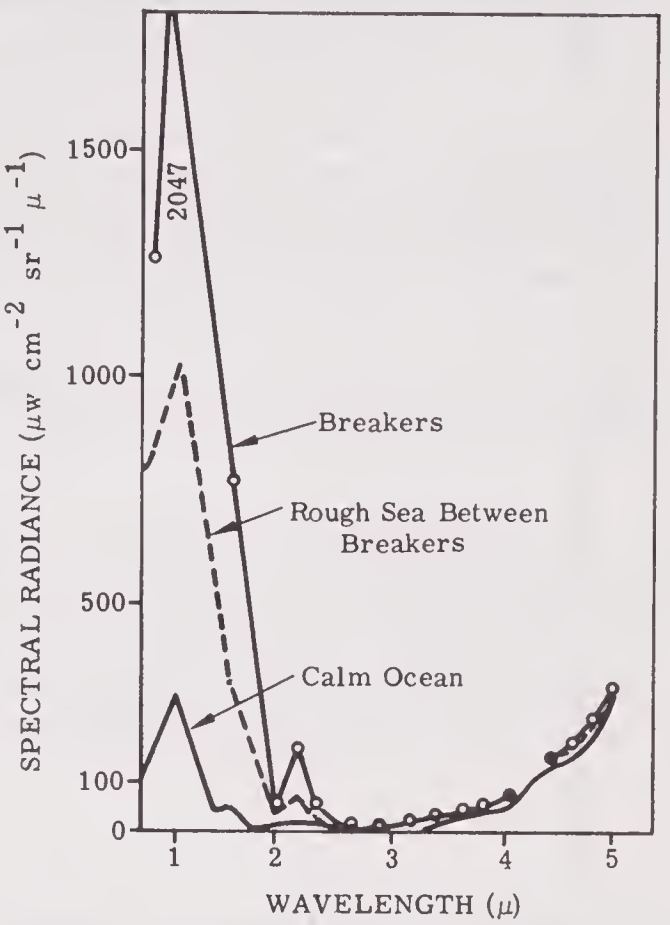


FIG. 5-96. Spectral radiance of the ocean [33].

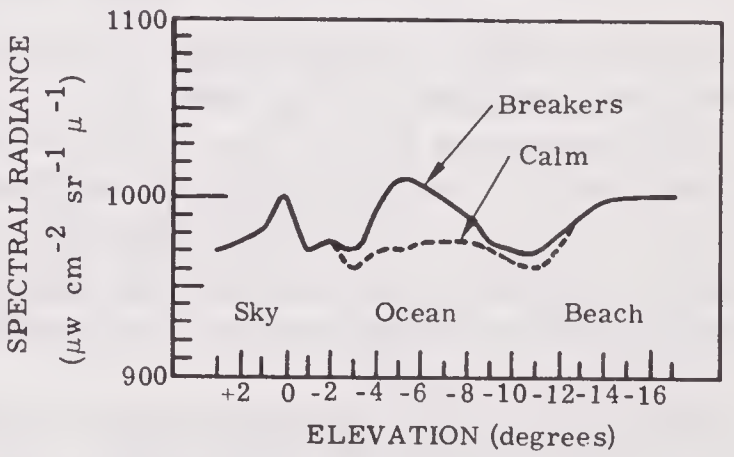


FIG. 5-97. Spectral radiance of the ocean *versus* elevation angle [33].

5.9.3. Sea-Surface Temperature Distribution. The temperature of the sea surface determines the contribution of emission to its total radiance. In arctic regions this temperature is near 0°C; near the equator it rises to 29°C. Currents such as the warm water of the Gulf Stream produce anomalies of several degrees centrigrade as it flows into colder areas. However, in most infrared scenes of marine interest, it is the radiance variation from point to point that determines the background against which a target is seen. Recent improvements in "thermal mappers" have shown details of this variation which is usually caused by temperature differences over the sea surface, but under some conditions reflected sky radiance predominates.

The temperature of the upper 0.1 mm of the sea surface under evaporative conditions has been measured as 0.6°C colder than water a few centimeters below [39]. The

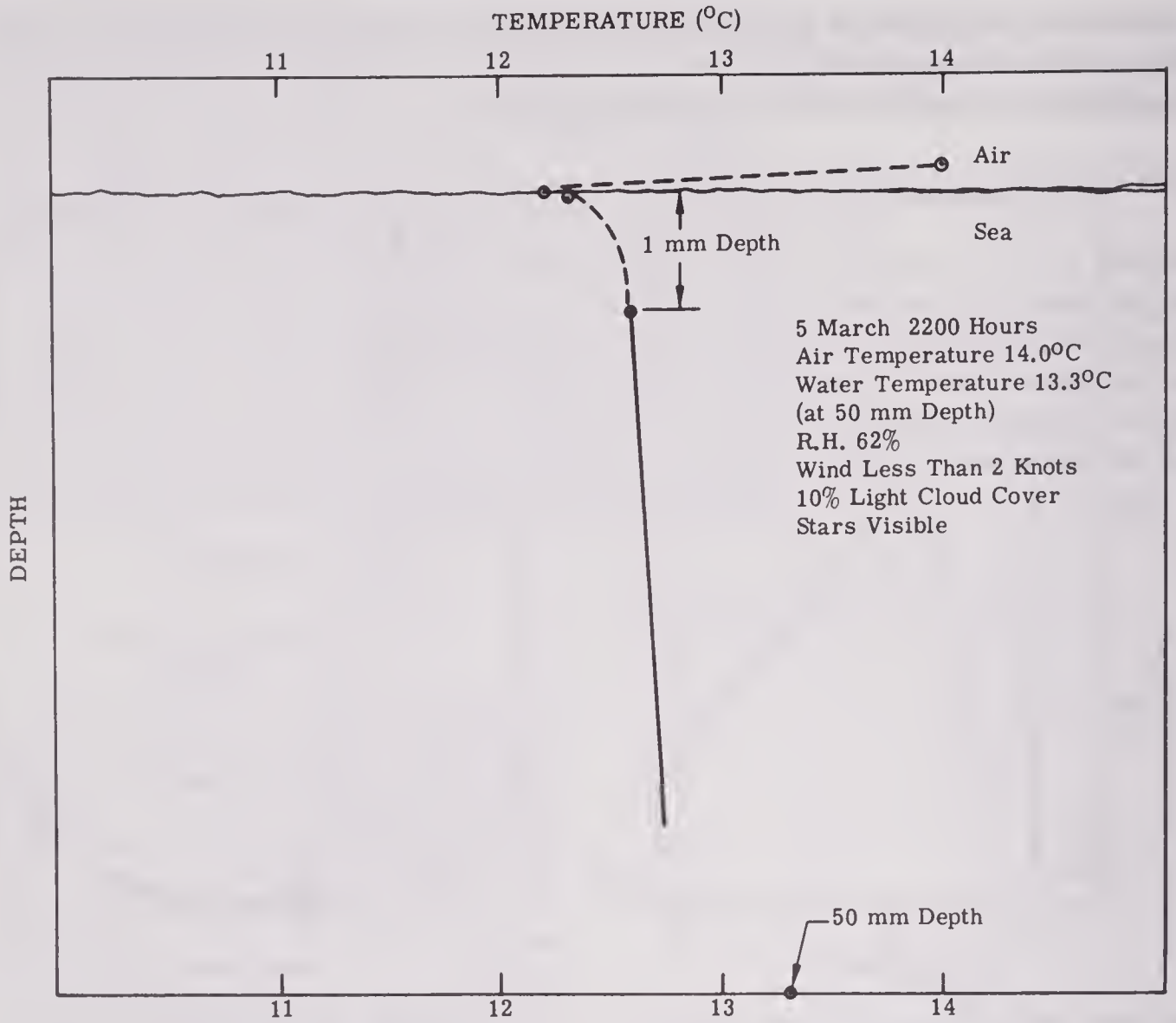


FIG. 5-98. Thermal structure of the sea boundary layer. Previous conditions: 12 hr cool (12° to 15°C), no rain. Data taken during passage of warm front.

sharpest gradient is in the upper 1 mm [40]. Measurements typical for the conditions noted are shown in Fig. 5-98.

The temperature of this layer with low heat capacity is determined by the rate of evaporation, by radiation exchange, and by the flow of heat from the air and from below. It has been found experimentally that the presence of surface contaminations reduces (slightly) the flow of heat from below so that a "slick" (a region in the sea with enough surface contamination to alter surface tension) appears colder than adjacent areas outside the slick.

Finally, the flow of heat from below is also influenced by the convective activity of the water layer above the thermocline.

5.9.4. Sky Radiance. For examples of sky radiance at night under clear, overcast, and other conditions refer to Sec. 5.2.

NOTES ADDED IN PROOF

5.1. Sky Backgrounds. There is a moderate amount of literature on the spatial and temporal fluctuations of the sky background [41-47].

5.2.1. Auroral Spectra. It is difficult to investigate the aurora and airglow beyond 2.0 μ because of absorption and thermal emission processes in the atmosphere. Reference [48] gives some predicted values for the 2.0-μ to 3.5-μ region. General reviews and one case of an application are covered in [49] through [52].

5.4.5. Spectral Distribution of Stellar Radiation. A useful source is [53], which includes reference [21] as an appendix. Additional work on infrared stellar sources has been done at Ohio State [54,55]. Reference [56] is a helpful catalog. Reference [57] updates the parameters used in converting from visual magnitudes to infrared spectral irradiance. Reference [58] is an example of an application. Other recent work is that of Low and Johnson [59] in the 10-μ to 20-μ region and of Leighton [60]. Reference [61] gives a 5-color statement of the magnitudes of 1300 bright stars.

One must at present use the visual magnitudes of bright stars to compute their infrared irradiance simply because the enormous job of cataloging the infrared emission from all the stars has hardly begun. The present approach will not adequately predict the irradiance from massive cool stars whose infrared magnitude might far exceed their visual magnitude [60].

5.5. The Earth as a Background. A comprehensive review of the unclassified literature on earth-background measurements taken from aircraft, satellites, rockets, and balloons is contained in [62].

5.8.2. Terrain Emissivity and Reflectivity. A complete catalog of spectral reflectance data of terrain from all available sources, reduced to a standard format of presentation, is now available [35]. The report itself is very bulky and has had limited circulation. However the data is on file at the Target Signatures Analysis Center, Willow Run Laboratories, at The University of Michigan's Institute of Science and Technology.

References

1. E. E. Bell, L. Eisner, J. Young, and R. A. Oetjen, *J. Opt. Soc. Am.*, **50**, pp. 1313-1320 (Dec. 1960).
2. E. E. Bell, I. L. Eisner, and R. A. Oetjen, "The Spectral Distribution of the Infrared Radiation from the Sky," *Proc. of the Symposium on Infrared Backgrounds,* Nonr-1224(12), Engineering Research Institute, University of Michigan (March 1956). AD 121010.
3. *Infrared Satellite Backgrounds,* Part I: *Atmospheric Radiative Processes,* AFCRL 1069(I), The Boeing Co., Aero-Space Div., Seattle, Wash. (Sept. 30, 1961).
4. J. W. Chamberlain, *Physics of the Aurora and Airglow,* (Academic Press, 1961).
5. A. W. Harrison and A. V. Jones, *J. Atmos. Terres. Phys.*, 11, pp. 192-199 (1957).
6. A. W. Harrison and A. V. Jones, *J. Atmos. Terres. Phys.*, **13**, pp. 291-294 (1957).
7. E. H. Vestine, *Terr. Magn.*, **49**, pp. 77-102 (June 1944).
8. F. W. G. White and M. Geddes, *Terr. Magn.*, **44**, pp. 367-377 (Dec. 1939).
9. L. Harang, *The Aurorae* (John Wiley & Sons, Inc., New York, 1951).
10. C. Störmer, *The Polar Aurora* (Clarendon Press, Oxford, 1955).
11. E. L. Krinov, "Spectra Reflectance Properties of Natural Formations." Originally published in Russian, 1947. Translation: National Research Council of Canada. Tech. Trans. TT439, Ottawa, Canada (1953).
12. F. F. Roach, *Proc. IRE,* **47**, p. 267 (1959).
13. J. F. Noxon, A. Harrison, and A. V. Jones, *J. Atmos. Terres. Phys.*, **16**, pp. 246-251 (1959).
14. A. Vallance Jones and H. Gush, *Nature,* **172**, p. 496 (Sept. 12, 1953).
15. D. M. Hunten, *J. Atmos. Terres. Phys.*, 7, pp. 141-151 (1955).
16. M. J. Seaton, *J. Atmos. Terres. Phys.*, **4**, pp. 285-294 (1954).
17. J. C. Brandt and J. W. Chamberlain, *J. Atmos. Terres. Phys.*, **13**, pp. 90-98 (Dec. 1958).
18. J. C. Brandt, *Astrophys. J.*, **128**, pp. 118-123 (1958).
19. J. C. Brandt, *Astrophys. J.*, **130**, pp. 228-240 (July 1959).
20. J. W. Chamberlain and C. Sagan, *Planet. Space Sci.*, **2**, pp. 157-164 (1960).
21. L. Larmore, *Infrared Radiation from Celestial Bodies,* U.S. Air Force Project, RAND Research Memo RM-793-1 (17 March 1952).
22. *Infrared Engineering Handbook,* Astrionics Div., Aerojet-General Corp. (25 Aug. 1961).
23. R. C. Ramsey, "Spectral Irradiance from Stars and Planets above the Atmosphere from 0.1 to 100.0 Microns," *Appl. Optics,* **VI**, 4, p. 465 (July 1962).
24. Jen-Hu-Chang, *Ground Temperature,* Blue Hill Meteorological Observatory (Harvard University Press, Cambridge, Mass., 1958).
25. F. A. Berry, et al., *Handbook of Meteorology* (McGraw-Hill Book Co., New York, 1945).

26. I. Solomon, *Estimates of Altitudes with Specified Probabilities of Being Above All Clouds,* Tech. Rept. 159, Air Weather Service (MATS), U.S. Air Force (Oct. 1961).
27. H. D. Nelgner and J. R. Thompson, *Airborne Spectral Radiance Measurements of Terrain and Clouds,* Emerson Electric Manufacturing Co., St. Louis, Mo., Rept. 1323, April 1962.
28. R. Schimpf and C. Aschenbrenner, *Z. Phot. Wiss. Tech.,* **2**, pp. 41-45 (1940).
29. Frank Benford, *Illum. Eng. Soc.,* **42**, pp. 527-544 (May 1947).
30. F. A. Brooks, *J. Meteorol.,* **9**, pp. 41-52 (1952).
31. W. R. Fredrickson, N. Ginsburg, and R. Paulson, *Infrared Spectral Emissivity of Terrain,* Final Report, Syracuse University Research Inst., Syracuse, N.Y. AD 155552.
32. D. M. Gates and W. Tantraporn, *Science,* **115**, pp. 613-616 (1952).
33. E. E. Bell, L. Eisner, J. Young, and R. A. Oetjen, *J. Opt. Soc. Am.,* **52**, pp. 201-209 (Feb. 1962).
34. W. R. Fredrickson, H. Ginsburg, R. Paulson, and D. L. Stierwalt, *Infrared Spectral Emissivity of Terrain,* Int. Dev. Rept. No. 2, AF33(616)-5034, Syracuse University Research Inst., Syracuse, N.Y. (Aug. 1, 1957).
35. *Target Signatures Study Interim Report,* Volume V: *Catalog of Spectral Reflectance Data,* The University of Michigan, Rept. No. 5698-22-T(V), October 1964.
36. E. D. McAlister, unpublished data, 1951-1952.
37. H. O. McMahon, *J. Opt. Soc. Am.,* **40,** pp. 376-380 (June 1950).
38. C. Cox and W. Munk, *Bull. Scripps Inst. Oceanog. Univ. Calif.,* **6,** pp. 401-488 (1956).
39. G. C. Ewing and E. D. McAlister, *Science,* **131,** pp. 1374-1376 (May 6, 1960).
40. E. D. McAlister, *J. Opt. Soc. Am.,* **52,** p. 607 (May 1962).
41. *Measurement of Infrared Radiation Gradients in the Sky,* Midwest Research Inst., Kansas City, Mo., 1953, AD 206 453.
42. R. C. Jones, *Sky Noise—Analysis of Circular Scanning,* Polaroid Corp., Cambridge, Mass., November 1953.
43. R. C. Jones, *Sky Noise—Its Nature and Analysis,* Polaroid Corp., Cambridge, Mass., September 1953.
44. R. E. Eisele, *Infrared Background Investigation,* Rept. AFCRC-TN59-843, Ramo-Wooldridge Division of Thompson-Ramo Wooldridge, Inc., Los Angeles, Calif., June 1959.
45. *Infrared Background Investigation,* Thompson-Ramo Wooldridge, Inc., Canoga Park, Calif., March 1960, AD 236 913.
46. H. E. Bennett, J. M. Bennett, and M. R. Nagel, *The Spatial Distribution of Infrared Radiation from the Clear Sky Including Sequences of Sky Maps at Various Elevations,* NAVORD Rept. 6577, U.S. Naval Ordnance Test Station, China Lake, Calif., September 1959.
47. H. E. Bennett, J. M. Bennett, and M. R. Nagel, *Measurements of Infrared and Total Radiance of the Clear Winter Sky at Wright-Patterson Air Force Base, Ohio,* Wright Air Development Center, USAF Air Research and Development Command, Wright-Patterson AFB, Ohio, March 1957, AD 118 127.
48. A. V. Jones, *Possible Methods for Studying the Auroral Spectrum in the 2.0 to 3.5 Micron Region,* Saskatchewan University, Saskatoon, November 1959.
49. D. M. Hunten, "Optics of the Upper Atmosphere," *J. Appl. Optics,* **3,** *2,* February 1964.
50. *Aurorae and Airglow,* National Aeronautics Space Administration, Washington, D.C., April 1964.
51. I. Sellin, *Auroral Radiations in the Infrared,* Laboratories for Applied Sciences, University of Chicago, October 1961.
52. R. Chapman, R. Jones, A. Dalgarno, and D. Beining, *Investigation of Auroral, Airglow and Night Emissions as Related to Space-Based Defense Systems,* Geophysics Corp. of America, Bedford, Mass., June 1962.
53. R. G. Walker, *Infrared Celestial Backgrounds,* Air Force Cambridge Research Labs., Bedford, Mass., July 1962.
54. P. E. Barnhart and W. E. Mitchell, Jr., *Stellar Background Measurement Program,* The Ohio State University, Columbus Research Foundation, Columbus, Ohio, July 1964.
55. *Space Background Study for Project DEFENDER,* Eastman Kodak Co., Rochester, N.Y., April 1963, AD 403 780.
56. *Space Handbook No. 1,* ACF Electronics Division ACF Industries Inc., Riverdale, Md., November 1962.
57. *Celestial Background Radiation,* Air Force Cambridge Research Labs., Bedford, Mass., March 1964, AD 602 616.
58. L. L. Collins and R. B. Freund, *Celestial Background Simulation Techniques,* Northrop Space Labs., Hawthorne, Calif., 1961, AD 282 788.
59. F. J. Low and H. L. Johnson, "Stellar Photometry at 10 μ," *J. Appl. Phys.* **139,** 1130 (1964).
60. R. Leighton, *Astrophys. J.* (Letters), No. 4 (1965).
61. "Magnitudes and Colors of 1300 Bright Stars," *Sky and Telescope,* **30,** *1,* 24 (July 1965).

62. N. Dittmar, F. Farley, and J. Boyse, *Earth Background Measurements: A Survey of the Unclassified Literature,* The University of Michigan, Willow Run Labs., Inst. of Science and Technology, Rept. No. 6054-16-X, in preparation.
63. A. Adel, *Observations of Atmospheric Scattering Near the Sun's Limb,* Arizona State College, Flagstaff, Ariz., January 1961, AD 273 599.
64. R. C. Jones and A. M. Nagvi, *Satellite Navigation by Terrestrial Occulations of Stars; III: Interference due to Brightness of the Earth's Atmosphere,* Geophysics Corp. of America, Boston, Mass., October 1962, AD 287 869.
65. A. M. Nagvi, *Satellite Navigation by Terrestrial Occulations of Stars—Considerations Relating to Refraction and Extinction,* Geophysics Corp. of America, Bedford, Mass., October 1962, AD 287 868.
66. G. Newkirk and J. Eddy, *Light Scattering by Particles in the Upper Atmosphere,* Colorado University, Boulder, May 1963.
67. R. K. McDonald, "High Altitude Sky Radiance Assessment," *Proc. IRIS,* **9,** *3,* Boeing Co., Seattle, Wash. (unclassified article in classified volume).
68. N. P. Laverty and W. M. Clark, "High-Altitude, Daytime Sky Radiance Measurements," *Proc. IRIS,* **9,** *3,* Boeing Co., Seattle, Wash. (unclassified article in classified volume).

Chapter 6
ATMOSPHERIC PHENOMENA

Gilbert N. Plass

Southwest Center for Advanced Studies

Harold Yates

Barnes Engineering Company

CONTENTS

6.1. Properties of the Atmosphere 177
6.1.1. Temperature 177
6.1.2. Pressure 178
6.1.3. Density 178
6.1.4. Atmospheric Composition 178
6.1.5. Particle Concentration and Size Distribution 187
6.2. Absorption by a Single Line 189
6.2.1. Single Line with Lorentz Shape 190
6.2.2. Single Line with Doppler Shape 191
6.2.3. Single Line with Both Doppler and Lorentz Broadening 192
6.3. Absorption by Bands 192
6.3.1. Elsasser Model 192
6.3.2. Statistical Model 194
6.3.3. Random Elsasser Model 196
6.3.4. Quasirandom Model 196
6.4. Useful Approximations to Band Models 197
6.4.1. Weak-Line Approximation 197
6.4.2. Strong-Line Approximation 200
6.4.3. Nonoverlapping-Line Approximation 201
6.5. Scattering 202
6.5.1. Relationship to Field of View 202
6.5.2. Meteorological Range 203
6.5.3. Scattering Coefficient 204
6.5.4. Scattering Coefficient Measurements 206
6.6. Atmospheric Scintillation 209
6.6.1. Inhomogeneities in the Atmosphere 210
6.6.2. Image Boil 212
6.6.3. Enlargement of the Image 213
6.6.4. Atmospheric Scintillation Measurements 214
6.7. Solar Spectrum Measurements 227

6.8. Total Absorption (Laboratory Measurements) 237
6.8.1. Total Absorption by CO_2 238
6.8.2. Total Absorption by H_2O 244
6.8.3. Total Absorption by N_2O 246
6.8.4. Total Absorption by CO 249
6.8.5. Total Absorption by CH_4 250
6.9. Infrared Transmission Through the Atmosphere 252
6.9.1. Horizontal Paths 252
6.9.2. Slant Paths 261
6.10. Calculation Procedures 266

6. Atmospheric Phenomena

6.1. Properties of the Atmosphere

6.1.1. Temperature. Standard-atmosphere temperature profiles from 0 to 100 km and from 0 to 700 km are shown in Fig. 6-1 and 6-2, respectively. The profiles are based on the U.S. Standard Atmosphere, 1962 [1].

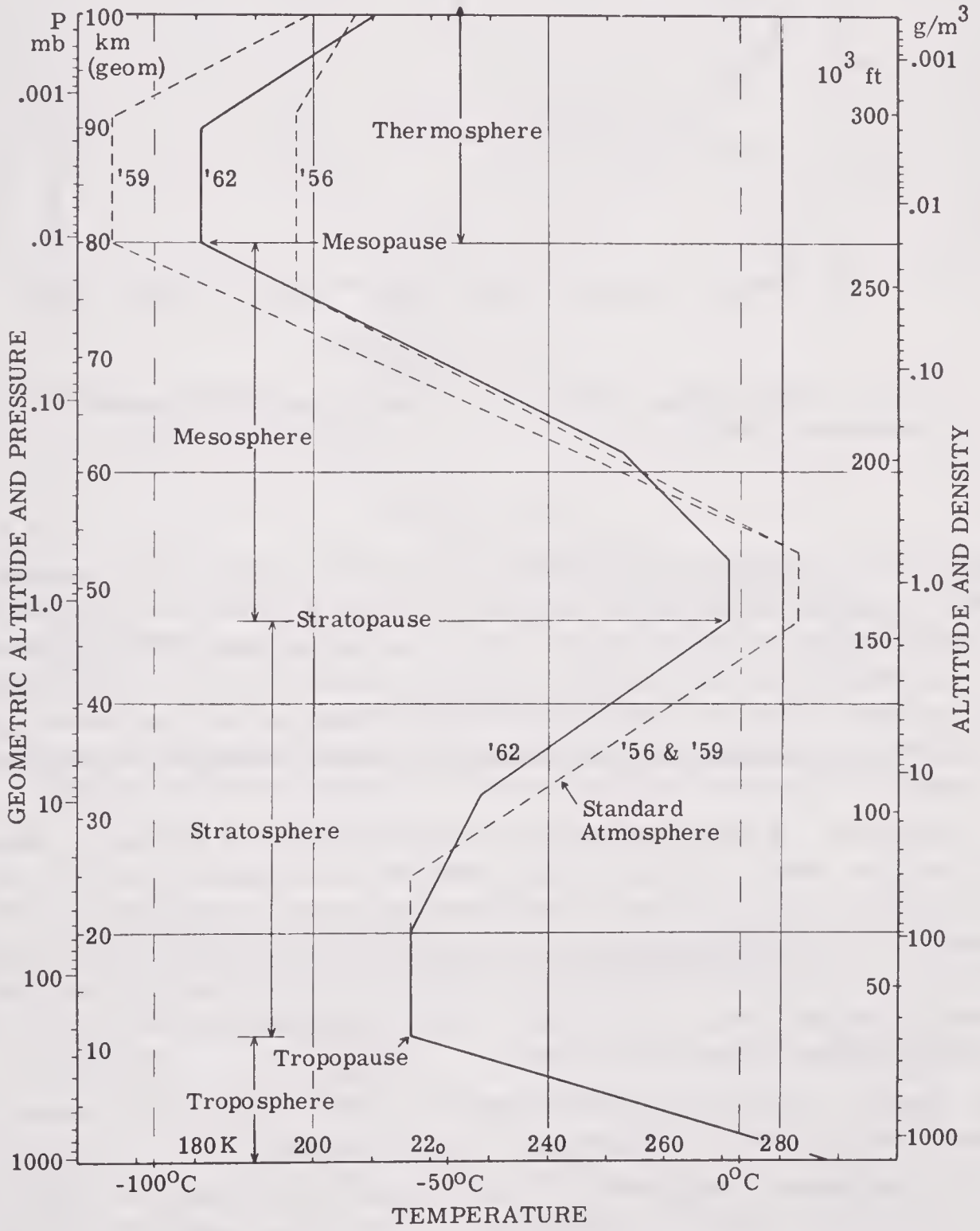


FIG. 6-1. Atmospheric temperature profiles from 0 to 100 km. Based on U.S. Standard Atmosphere, 1962 [1].

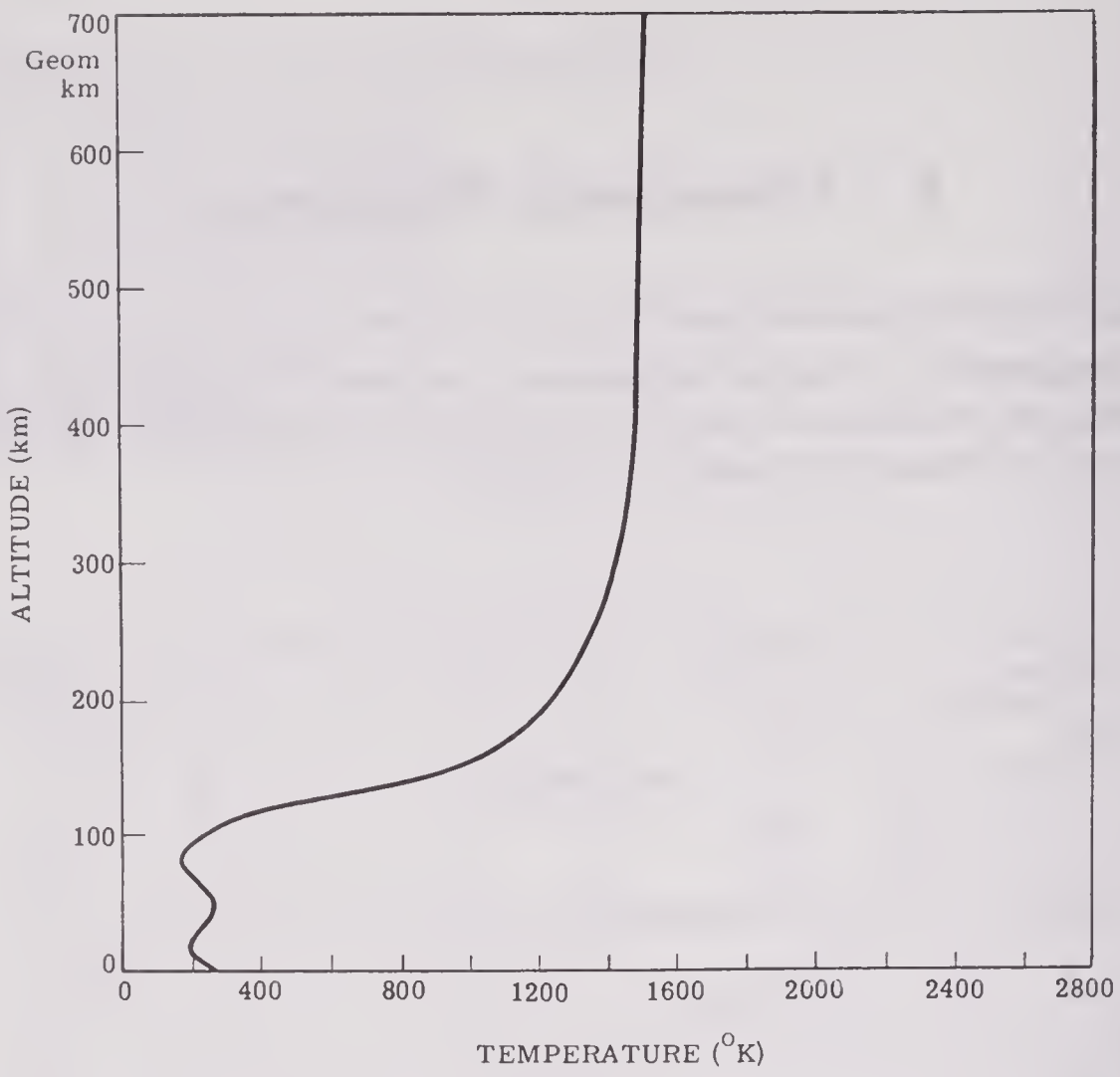


FIG. 6-2. Temperature profile from 0 to 700 km. Based on U.S. Standard Atmosphere, 1962 [1].

Supplemental atmospheric temperature profiles [2], which reflect seasonal and latitudinal variability up to 90 km, are shown in Fig. 6-3 and 6-4. Winter profiles typical of the tropics (15°N), subtropics (30°N), and midlatitudes (45°N) are shown in Fig. 6-3; summer profiles for the same areas are shown in Fig. 6-4. The winter and summer profiles for 15°N are identical and actually represent a mean annual profile; at this latitude, the temperature-height structure remains relatively constant throughout the year.

6.1.2. Pressure. Standard-atmosphere pressure from 0 to 100 km is shown in Fig. 6-1; the left ordinate of Fig. 6-1 lists pressure versus height from 0 to 100 km. All of the pressure data are based on the United States Revised Atmosphere, 1962 [1].

Supplemental atmospheric-pressure data, which reflect seasonal and latitudinal variability (Sec. 6.1.1), are given in [2]. The differences in pressure between the supplemental atmospheres, and between each supplemental atmosphere and the standard atmosphere (Fig. 6-1), are slight.

6.1.3. Density. Standard-atmospheric-pressure data, which reflect seasonal and latitudinal variability (Sec. 6.1.1), are given in [2] and Fig. 6-1.

6.1.4. Atmospheric Composition. Table 6-1 gives the composition of the atmosphere up to about 90 km. Only carbon dioxide (CO_2), water vapor (H_2O), ozone (O_3), methane (CH_4), nitrous oxide (N_2O), and carbon monoxide (CO), are discussed in this section. The two most abundant gases, N_2 and O_2, although they do not exhibit any infrared absorption bands, affect the intensities of the observed absorption bands of the other constituents through Lorentz (pressure, collision) broadening (Sec. 6.2).

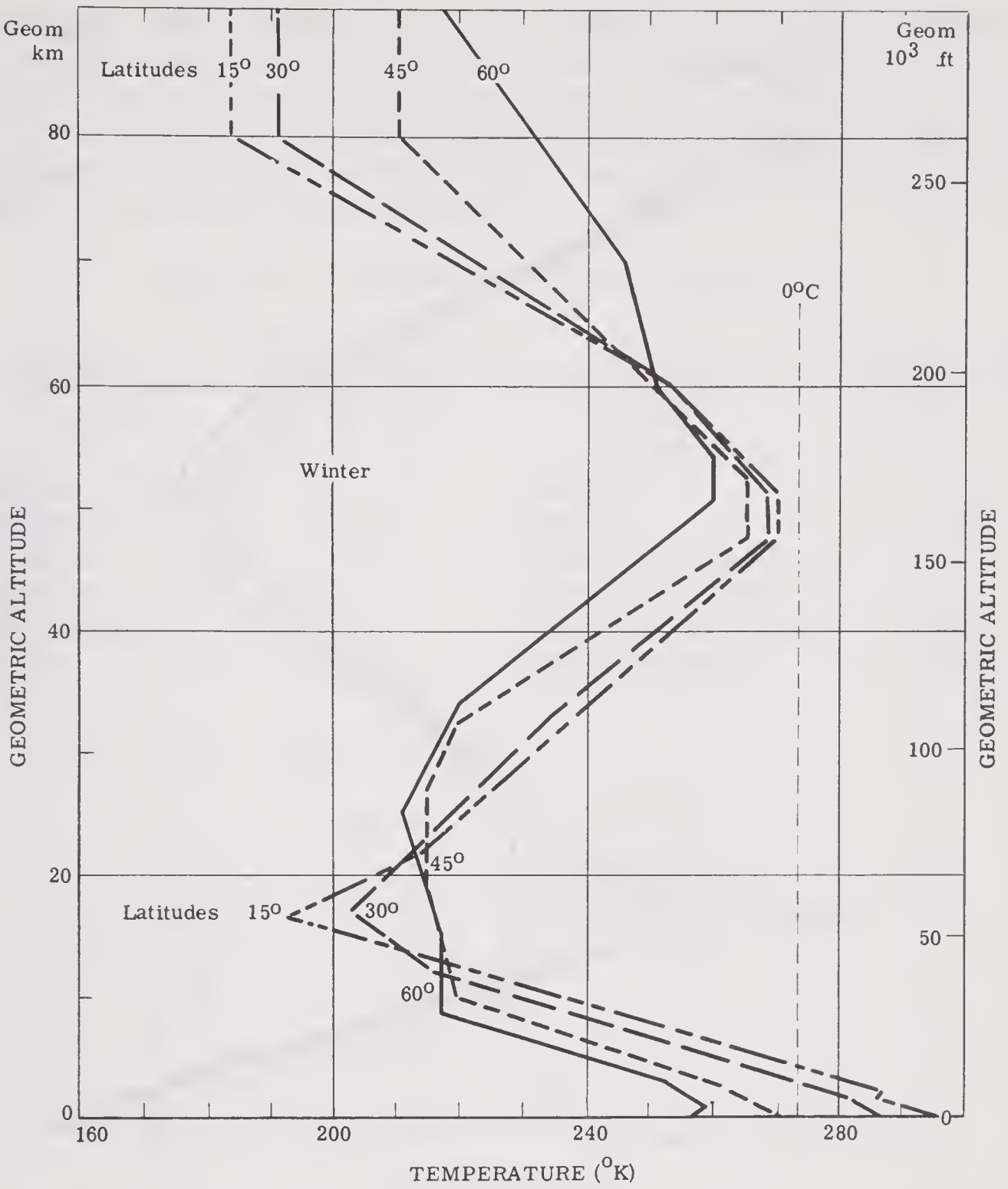


FIG. 6-3. Supplemental winter temperature profiles for tropics (15°N), subtropics (30°N), mid-latitudes (45°N), and subarctic latitudes (60°N) [2].

TABLE 6-1. COMPOSITION OF THE ATMOSPHERE [3].

Constituent	*Percent by Volume*	*Constituent*	*Percent by Volume*
Nitrogen	78.088	Krypton	1.14×10^{-4}
Oxygen	20.949	Nitrous Oxide	5×10^{-5}
Argon	0.93	Carbon Monoxide	20×10^{-6}
Carbon Dioxide	0.033	Xenon	8.6×10^{-6}
Neon	1.8×10^{-3}	Hydrogen	5×10^{-6}
Helium	5.24×10^{-4}	Ozone	variable
Methane	1.4×10^{-4}	Water Vapor	variable

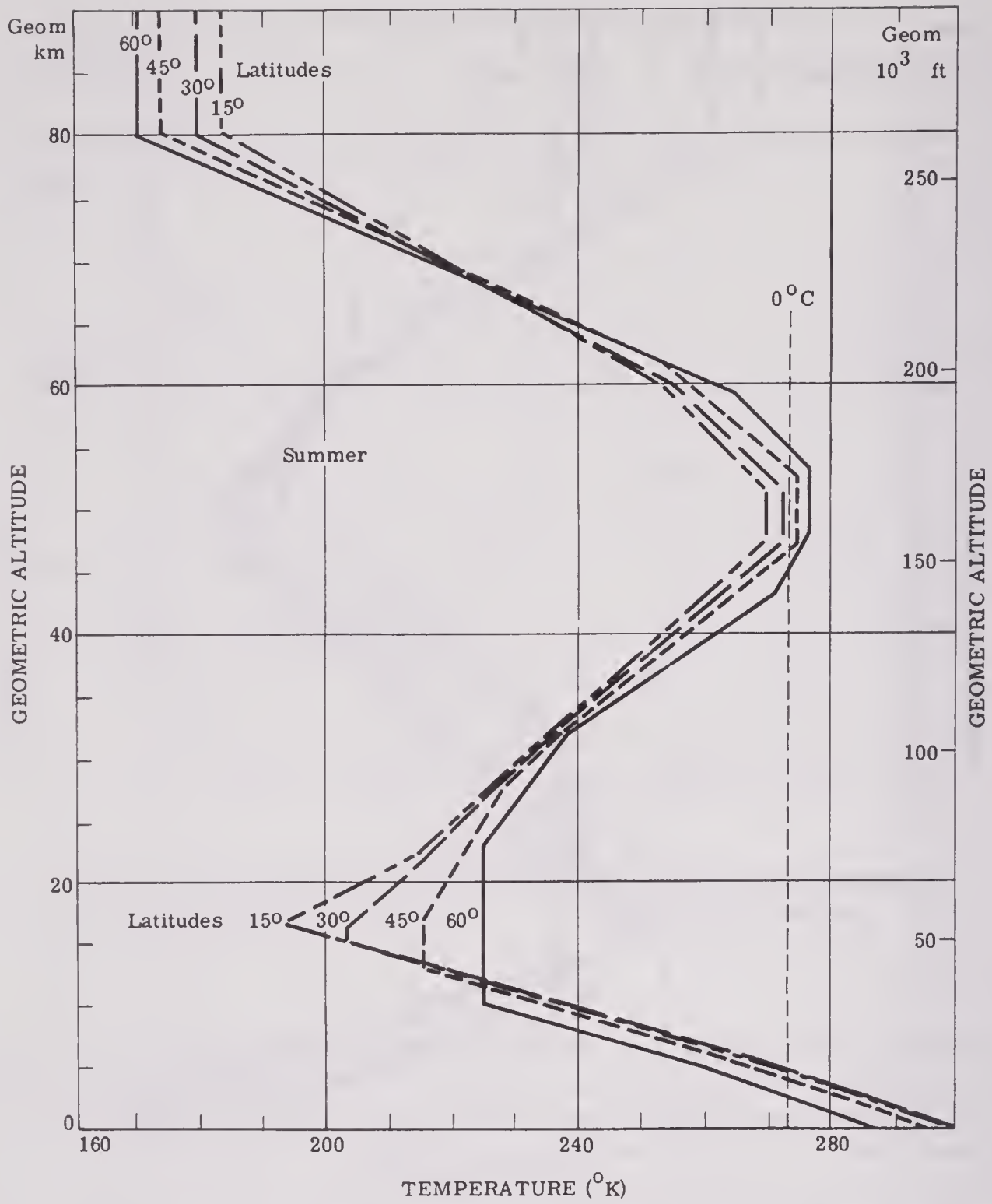


FIG. 6-4. Supplemental summer temperature profiles for tropics (15°N), subtropics (30°N), mid-latitudes (45°N), and subarctic latitudes (60°N) [2].

6.1.4.1. *Carbon Dioxide Distribution.* The average amount of CO_2 present in the atmosphere is 0.33% by volume. This average concentration remains almost constant in both space and time (Table 6-2). The burning of fossil fuels gives rise to a gradual increase of about 0.7 ppm per year. Superimposed on this fundamental cycle is a ground-level daily cycle caused by the exchange of CO_2 with soil and vegetation. The ground-level daily cycle can have local fluctuations of CO_2 concentration ranging from 200 to 600 ppm, but is characteristically confined to a shallow layer of the atmosphere immediately above the earth's surface (up to a few hundred feet). In addition, there is a seasonal variation, apparently due to vegetation, which depends on latitude and is about 2 ppm between 45° and 90°N.

In view of both the small departure in CO_2 concentration over the longer cyclic period and the uniformity of horizontal CO_2 distribution, it may be assumed that no noticeable variations of the mixing ratio occur with height above the biologically active

ground layer. Thus, the annual mean CO_2 concentration profile shown in Fig. 6-5 is reasonably accurate, at least above a few hundred feet altitude.

Below this altitude, data on the distribution of CO_2 over a particular area at a particular time for absorption and transmission calculations is desirable. In most cases, however, such data are not available and the mean profile will probably have to used, with some degradation in accuracy expected.

TABLE 6-2. VARIATION OF CO_2 CONCENTRATION [4].

Air Mass Time of Year Location	*Average Variation from Average Concentration (%)*	*Air Mass Time of Year Location*	*Average Variation from Average Concentration (%)*
Maritime air (Europe)	−1.1	Autumn	−0.1
Continental air	+1.6	Winter	+0.6
Polar air	−0.7	Rural France (64°N)	+4.4
Tropical air	+2.5	West Indies (20°N)	+3.3
Spring (Europe)	+0.8	South America (40°S)	−1.0
Summer	−1.4	Cape Horn (56°S)	−6.9

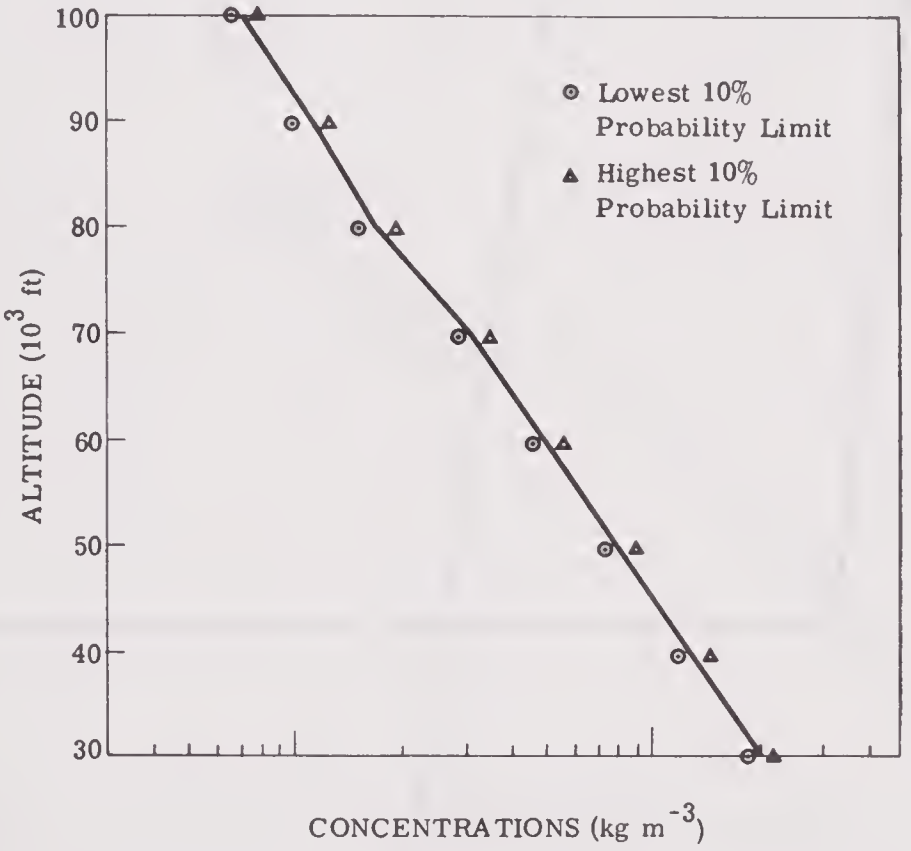


FIG. 6-5. Annual mean CO_2 concentration *vs.* altitude.

6.1.4.2. *Water-Vapor Distribution.* Water-vapor measurements are generally in agreement below the tropopause. In the stratosphere, however, there is wide disagreement among various measurements [5]. Some indicate a relatively dry atmosphere with a constant mixing ratio of about 0.05 g/kg (dry stratosphere); other measurements indicate that there is a recovery of the mixing ratio from about 0.002 or 0.003 at the tropopause to about 0.1 near 30 km (wet stratosphere).

There is no agreement at the present time between those who favor a "dry" stratosphere [6-16] and those who favor a "wet" stratosphere [17-24]. Accordingly, water-vapor distribution for a dry stratosphere is shown in Fig. 6-6 and for a wet stratosphere in Fig. 6-7 and 6-8.

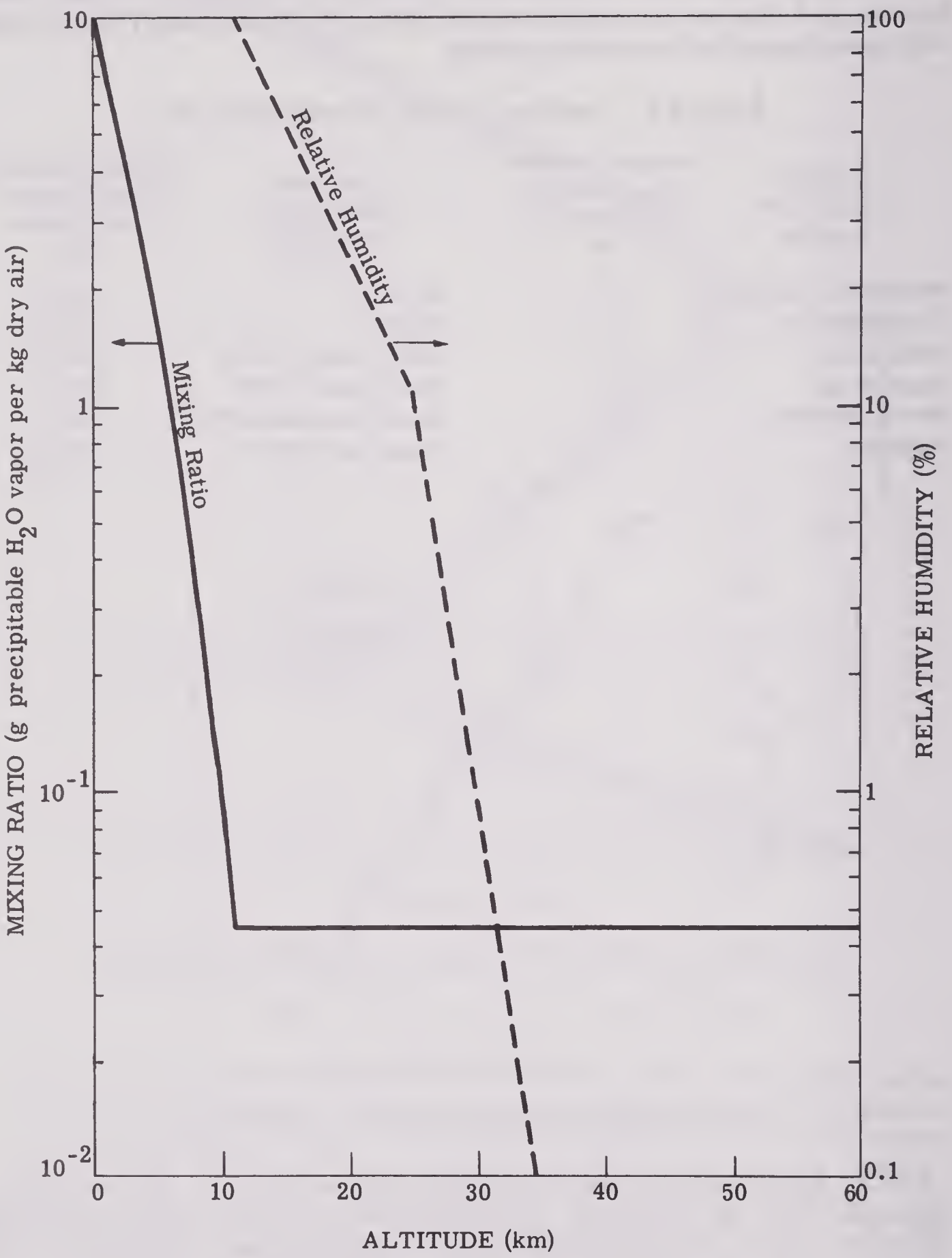
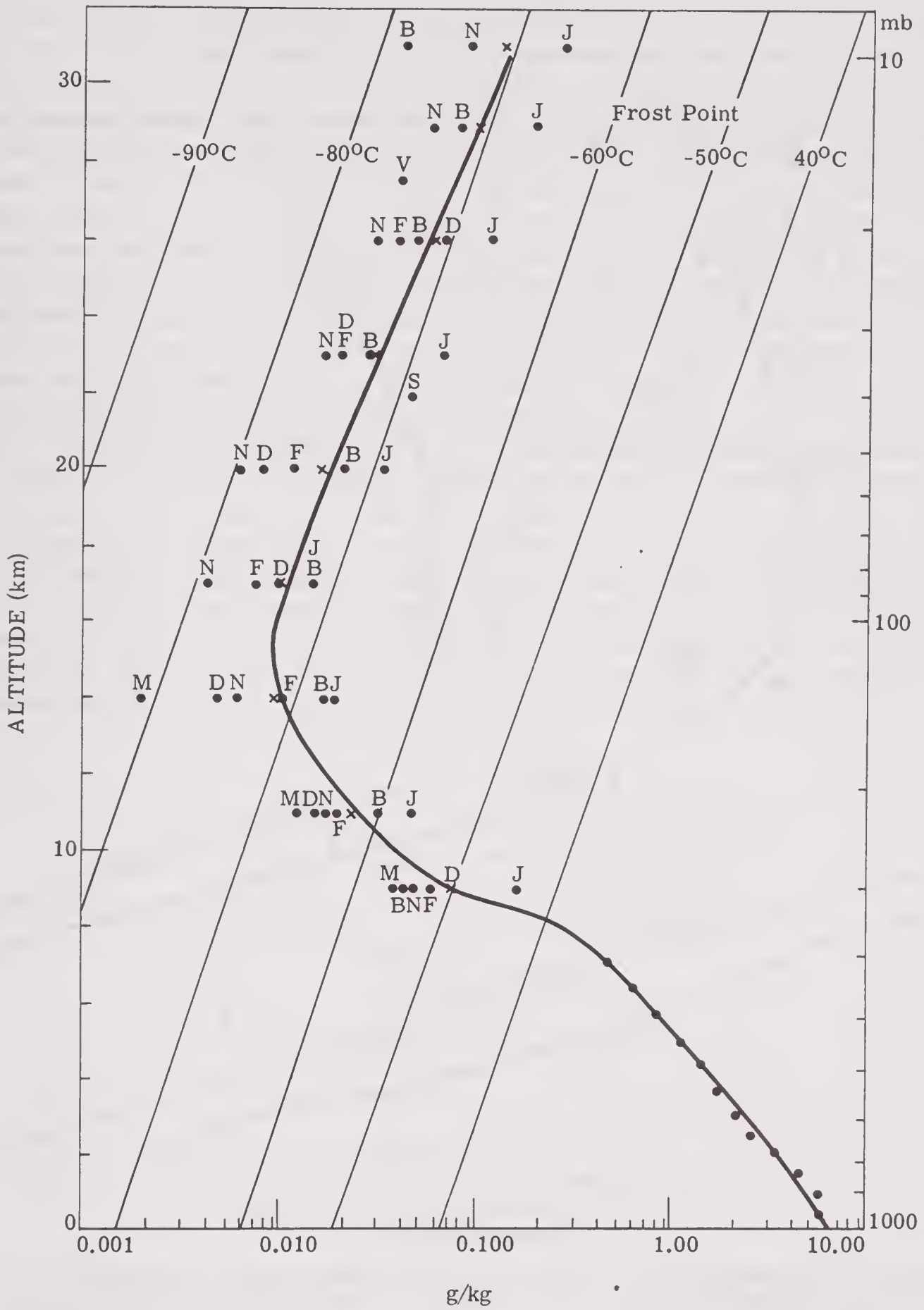


FIG. 6-6. Water vapor mixing ratio *vs.* altitude ("dry" stratosphere) [25].



N - Mean of 2 hydrometric ascents by U. S. Naval Research Labs.

B - Mean of 2 hydrometric ascents by Ballistic Research Labs.

F - One spectroscopic ascent by Univ. of Denver

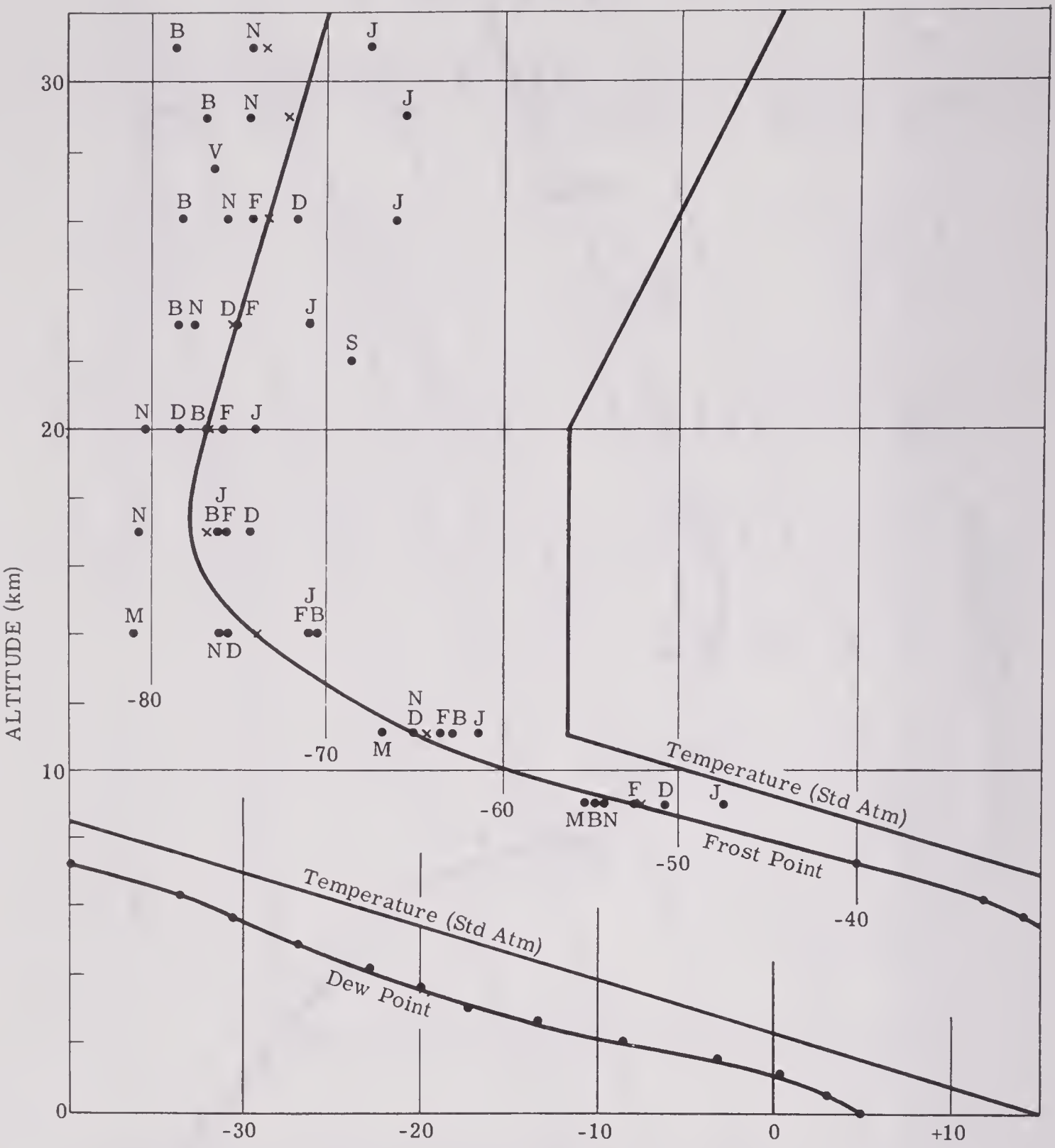
J - Means of the Japanese Meteorological Agency ascents, 100 of which reached or exceeded 300 mb and 2 of which reached 10 mb

M - Means of the British Meteorological Research Flight hydrometric ascents, 400 of which reached or exceeded 300 mb

V - Mean of 7 ascents of the United Kingdom Atomic Energy Authority's water vapor absorption device (vapor trap)

X - Mean of all symbols at level

FIG. 6-7. Water vapor mixing ratio *vs.* altitude ("wet" stratosphere) [26].



TEMPERATURE (°C)

N - Mean of 2 hydrometric ascents by U. S. Naval Research Labs.

B - Mean of 2 hydrometric ascents by Ballistic Research Labs.

F - One spectroscopic ascent by Univ. of Denver

D - One hydrometric ascent by Univ. of Denver

J - Means of the Japanese Meteorological Agency ascents, 100 of which reached or exceeded 300 mb and 2 of which reached 10 mb

M - Means of the British Meteorological Research Flight hydrometric ascents, 400 of which reached or exceeded 300 mb.

V - Mean of 7 ascents of the United Kingdom Atomic Energy Authority's water vapor absorption device (vapor trap)

X - Mean of all symbols at level

FIG. 6-8. Water vapor dewpoint and frost point *vs.* altitude ("wet" stratosphere) [26].

6.1.4.3. *Ozone Distribution* [24]. Of the ozone in the atmosphere 90% is concentrated in a layer between about 10 and 30 km above the earth's surface. The maximum concentration occurs between about 25 and 30 km.

In the upper atmosphere, ozone is formed by the photochemical dissociation of oxygen caused by radiation at wavelengths shorter than 2530 A. The oxygen atoms subsequently combine with an oxygen molecule to form ozone. In the lower atmosphere, minor amounts of ozone are thought to be formed through photochemical reduction of atmospheric pollutants. Ultraviolet radiation between 2000 and 2900 A breaks down ozone, resulting in an equilibrium in ozone formation and destruction. Should an imbalance in ozone concentration at any particular altitude occur, the time required to restore equilibrium can be determined. At altitudes of about 50 km, this time is a matter of minutes, below 35 km a matter of days, and below 15 km a matter of years.

Low-altitude ozone concentrations are influenced by atmospheric motions, dust, and other attenuators that interfere with the establishment of equilibrium conditions [4, 27-29]. For example, a large-scale flattening out causes a low-altitude secondary maximum in the ozone profile near the tropopause. When such conditions occur, photochemical processes quickly restore equilibrium in the higher altitudes and provide an overall increase in the total ozone content of the atmosphere. Hence, as shown in Fig. 6-9, a correlation exists between the maximum level of ozone and the total amount of ozone [30]. Two levels of maximum concentration are indicated: (1) a high-altitude (27 km) maximum of about 10 to 15 $\times$ 10^3 cm km^{-1} that occurs during all seasons, and (2) a low-altitude maximum (12 km) that occurs only during the winter and only at times of high total ozone concentration.

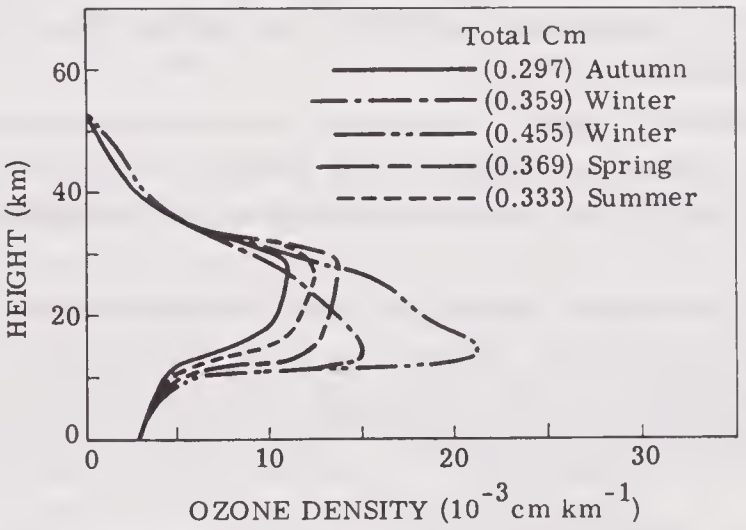


FIG. 6-9. Mean seasonal vertical distribution of ozone (approximately 52°N) [29].

Table 6-3, based on limited data, presents probable ozone distributions. The large increase at the 12-24-km altitude reflects the winter correlation between total ozone and lower stratospheric temperature changes (sudden warnings).

TABLE 6-3. VERTICAL DISTRIBUTION OF OZONE FOR TWO TOTAL OZONE CONCENTRATIONS [31].

Total Ozone (*cm*)	*Altitude* (*km*)			
	0–12	12–24	24–36	36–54
0.300	15%	37%	37%	11%
0.500	12%	53%	25%	8%

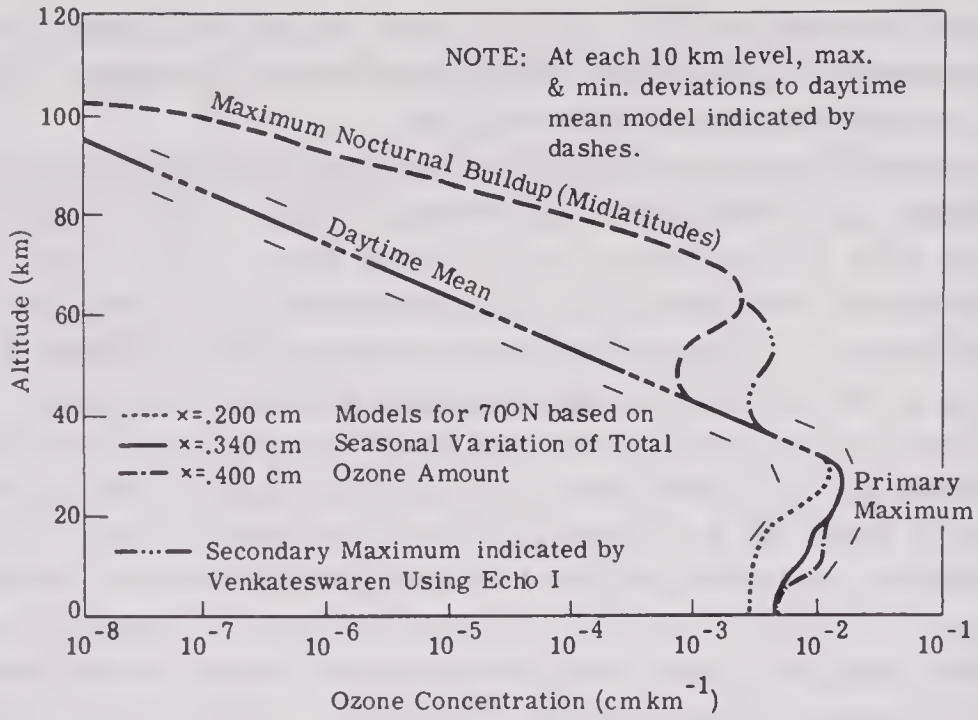


FIG. 6-10. Mean ozone distribution [24].

Figure 6-10 shows the mean ozone distribution for the Northern Hemisphere. Figure 6-11 and 6-12 show the seasonal variation of total ozone and the mean meridional distributions of ozone, respectively. The curves in Fig. 6-10 are essentially an envelope of possible distributions within which it is presumed all distributions will occur. The primary maximum concentration shown at an altitude of 25 to 30 km is well founded on both theoretical and experimental data. The remainder of the profile, however, is established in part by theoretical considerations of the effect of the normal process of weather on ozone distribution [27].

Figure 6-10 also shows suggested maximum and minimum concentration extremes for each 10-km level based on highest and lowest concentration values on record. These minimum and maximum extremes should not be interpreted as minimum and maximum profiles.

The nocturnal high-altitude buildup of ozone concentrations shown in Fig. 6-10 is, to a great extent, conjecture. Theory predicts that such a secondary maximum should occur at about 70 km at night, primarily as a result of the large concentrations of atomic oxygen at these heights [32, 34]. Some evidence of the existence of such a

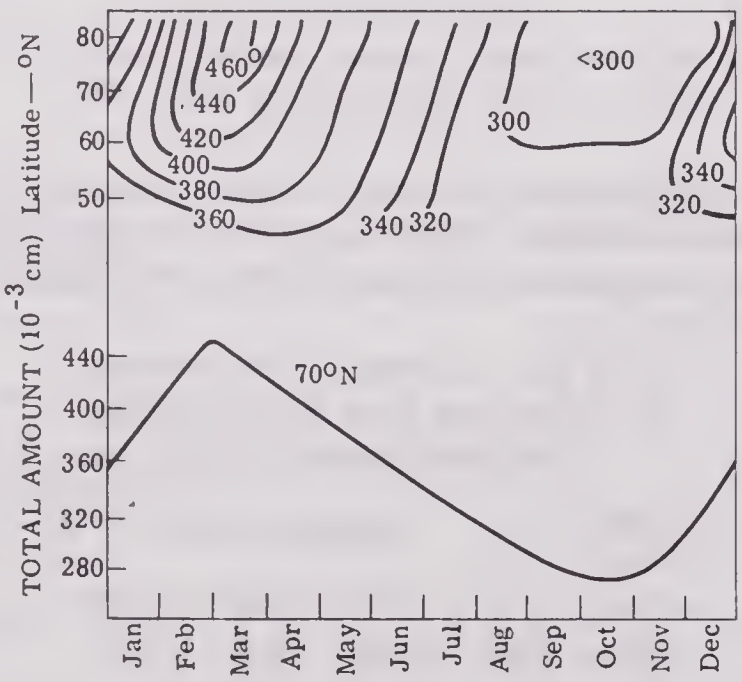


FIG. 6-11. Seasonal variation of total ozone [32].

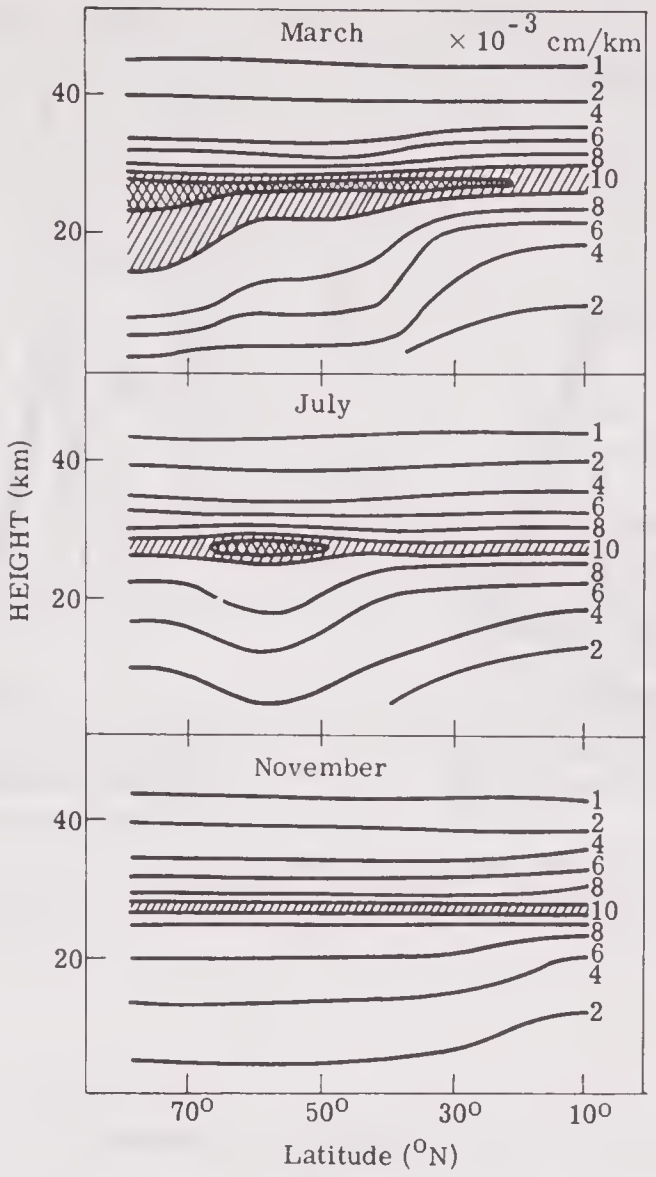


FIG. 6-12. Mean meridional distribution of ozone [33].

high-altitude maximum is presented by spectral observations of Echo I brightness [35]. In this instance, the high-altitude maximum occurs at about 55 km, and can be explained only partially by photochemical theory. System inaccuracies and questionable assumptions may account for some of the observed results.

6.1.4.4. *Methane, Nitrous Oxide, and Carbon Monoxide Distributions.* Methane, nitrous oxide, and carbon monoxide are rarer gases that absorb significantly only over long paths. The amount of these gases in the atmosphere is given in Table 6-1. All available evidence indicates a uniform mixing ratio [36].

6.1.5. Particle Concentration and Size Distribution. The concentration and size distribution of scattering particles in the atmosphere vary widely both geographically, and temporally in a given location. Accurate measurement of these particles is difficult, and the amount of good data is limited. From the available data, however, representative samples of the distribution of particles with respect to size are plotted in Fig. 6-13 through 6-16. Figure 6-13 gives an average particle-size distribution curve for a continental air condition and for a maritime air condition [37]. Figure 6-14 is the distribution of particles measured by capture for a haze and a fog [38]. Figure 6-15 is the relative size-distribution curve for a fair-weather cumulus cloud, where the total number of particles per cm^3 in the cloud is 300 [39]. So that measured values may be compared with analytical functions used to approximate the size-distribution curves of natural aerosols, Fig. 6-16 shows three such functions, two haze models and one cloud model [40]. The curves are normalized so that the integrated area under each curve gives 100 particles per cm^3. The concentration of scattering particles as a function of altitude is shown in Fig. 6-17, which includes the results of a number of investigators [41].

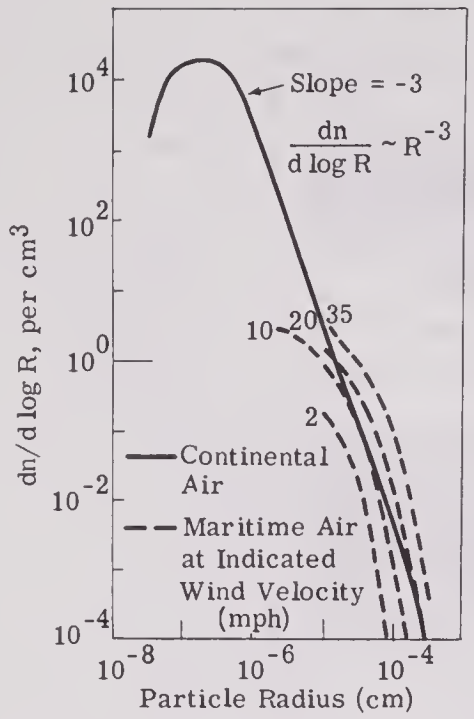


FIG. 6-13. Particle-size distribution for continental and maritime air [37].

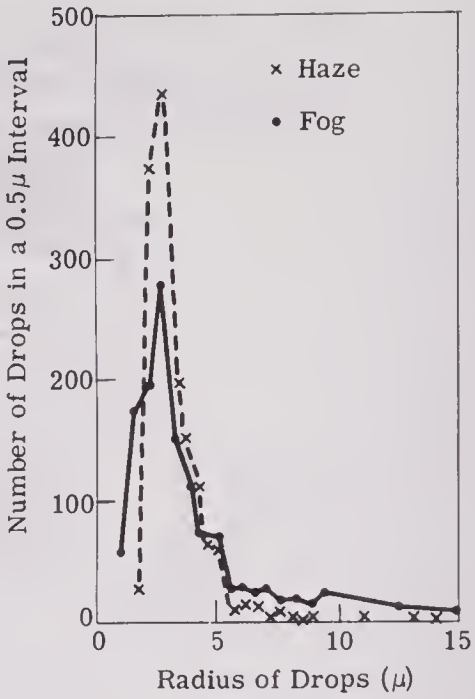


FIG. 6-14. Particle size distribution measured by capture for a haze and a fog [38].

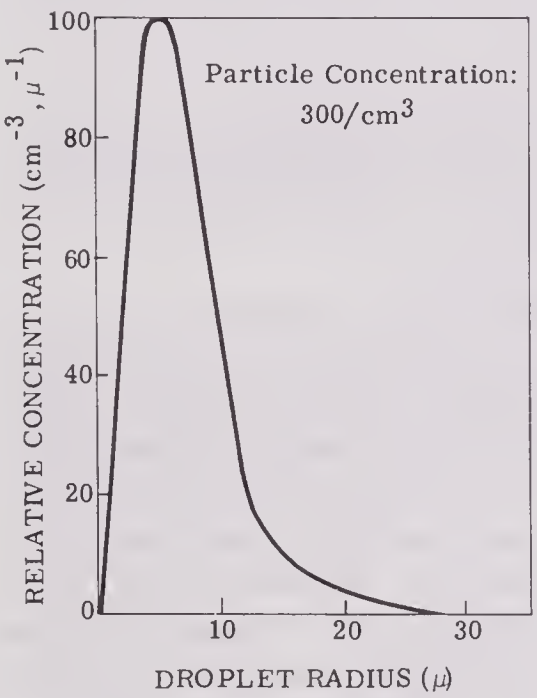


FIG. 6-15. Relative size-distribution of droplets in a fair-weather cumulus cloud [39].

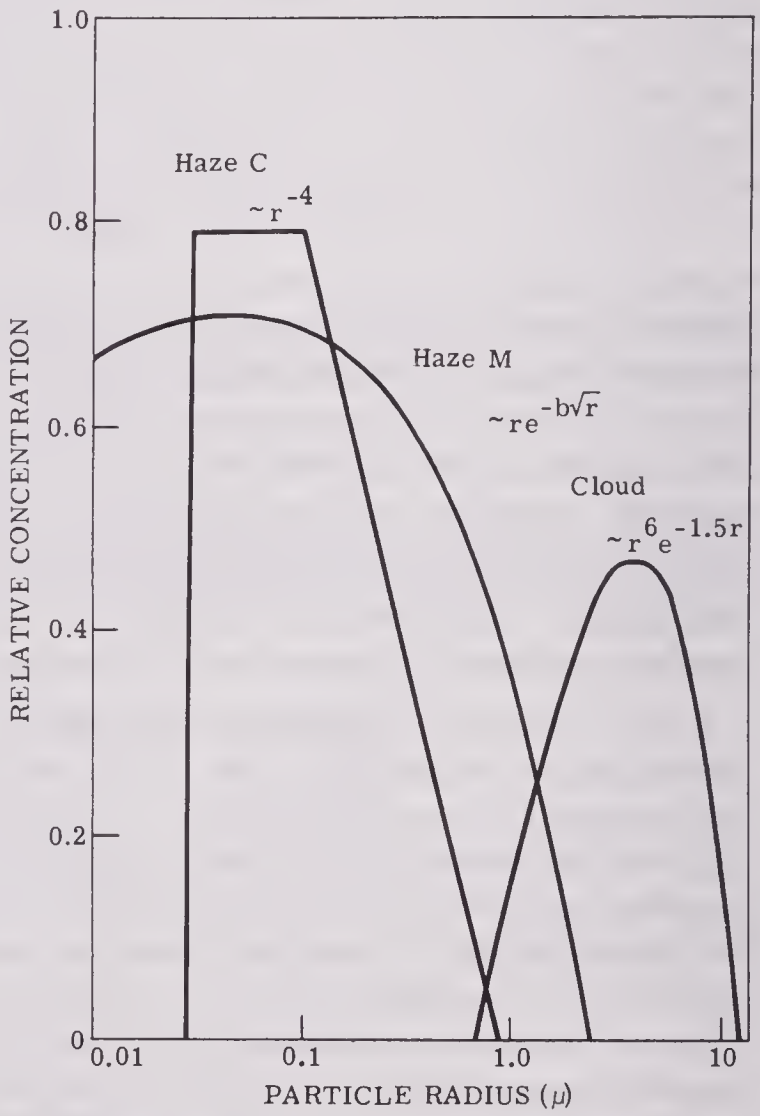


FIG. 6-16. Three model size-distribution functions [40].

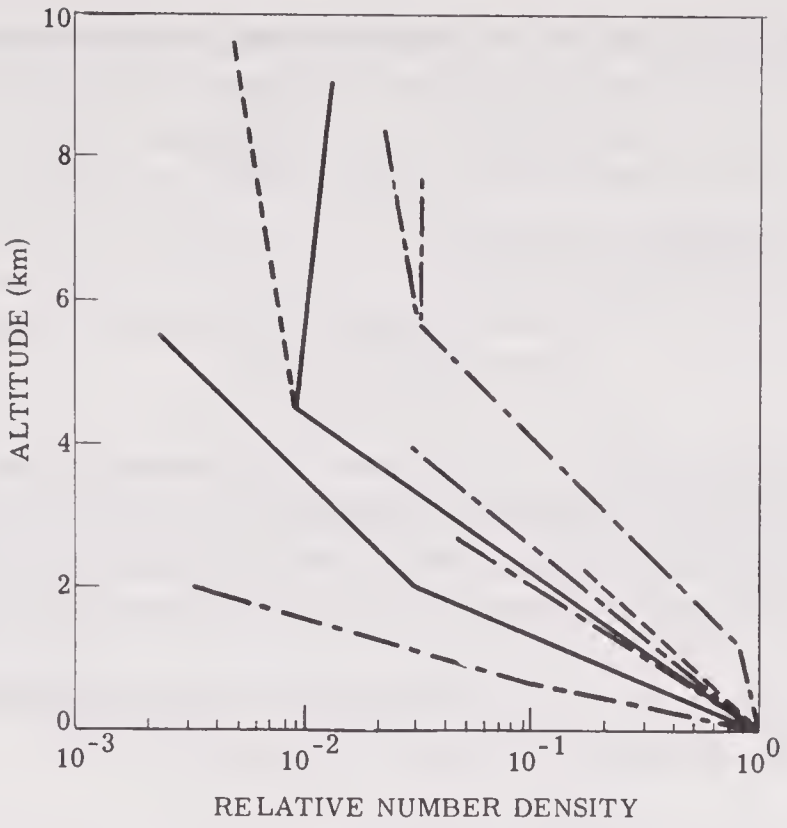


FIG. 6-17. Relative particle concentration as a function of altitude from various investigators [41].

6.2. Absorption by a Single Line

The fractional radiant absorptance, A, over a finite wave-number interval, $\Delta\nu$, is defined as

$$A\,\Delta\nu = \int_{\Delta\nu} (1 - e^{-k_\nu u})\, d\nu \tag{6-1}$$

where k_ν is the absorption coefficient at the wave number and u is the mass of absorbing gas per unit area. For most atmospheric absorption problems, the Lorentz, or pressure-broadened, line shape should be used. At very great altitudes in the stratosphere, the principal cause of line broadening is the Doppler effect.

The absorption coefficient, k_ν, for the Lorentz line shape is

$$k_\nu = \frac{S}{\pi} \frac{\alpha}{(\nu - \nu_0)^2 + \alpha^2} \tag{6-2}$$

where S is the total line intensity and α is the half-width of the spectral line whose center is located at the wave number ν_0. According to kinetic theory, the half-width depends on both the pressure, p, and absolute temperature, T, as

$$\alpha = \alpha_0 \left(\frac{p}{p_0}\right) \left(\frac{T_0}{T}\right)^{1/2} \tag{6-3}$$

where 0 refers to the value of the quantity for some standard condition. For most atmospheric infrared problems, pressure, p, can be taken as the total pressure. It is more accurate to use an effective pressure equal to the total pressure plus some constant times the partial pressure [42]. If the absorbing gas is only a small fraction of the total, however, the difference between the two quantities is usually small.

When line broadening is due to the Doppler effect, the absorption coefficient is [43]

$$k_\nu = \frac{S}{\Delta\nu_D}\left(\frac{\ln 2}{\pi}\right)^{1/2} \exp\left[-\frac{\ln 2}{\Delta\nu_D{}^2}(\nu-\nu_0)^2\right] \tag{6-4}$$

where the Doppler half-width is given by

$$\Delta\nu_D = \frac{\nu_0}{c}\left(\frac{2kT}{m}\ln 2\right)^{1/2} \tag{6-5}$$

and k is the Boltzmann constant, m is the mass of the molecule, and c is the velocity of light. Although the Lorentz line shape decreases only as $(\nu-\nu_0)^2$ at wave numbers that are more than several half-widths from the line center, the Doppler line shape is concentrated more near the line center and falls off exponentially in the wings of the line.

When both the processes that lead to the Lorentz and Doppler line shapes must be considered at once, the absorption coefficient is given by [43, 44]

$$k_\nu = \frac{(\ln 2)^{1/2}}{\pi^{3/2}}\frac{Sa}{\Delta\nu_D}\int_{-\infty}^{\infty}\frac{e^{-x^2}}{a^2+(w-x)^2}dx \tag{6-6}$$

where

$$a = (\ln 2)^{1/2}\frac{\alpha}{\Delta\nu_D} \tag{6-7}$$

$$w = (\ln 2)^{1/2}\frac{\nu-\nu_0}{\Delta\nu_D} \tag{6-8}$$

$$x = Su/2\pi\alpha \tag{6-9}$$

Although this integral cannot be evaluated in closed form, numerous approximations are given in the literature. A general expression for the Taylor series and the asymptotic expression is given in [44]; [43] gives a general review of the problem.

6.2.1. Single Line with Lorentz Shape. When the pressure-broadened Lorentz line shape is valid, the absorptance as obtained from Eq. (6-1 and 6-2) can be written as

$$A\,\Delta\nu = \alpha\int_{-\infty}^{\infty}\left[1-\exp\left(\frac{-2x}{1+\nu'^2}\right)\right]d\nu' \tag{6-10}$$

where

$$\nu' = (\nu-\nu_0)/\alpha \tag{6-11}$$

The limits of integration have been extended to infinity since it is assumed that there is no absorption outside of the interval $\Delta\nu$.

Two limiting results for the absorptance that can frequently be used are the weak-line and strong-line approximations (Sec. 6.4.1 and 6.4.2). When the path length is small or the pressure is large, the absorption is small at all wave numbers, including the line centers. In this case the exponential in Eq. (6-10) can be replaced by the first two terms in its series expansion, and the fractional radiant absorption, A, over a finite wave-number interval, $\Delta\nu$, can be expressed as [45-47]

$$A\,\Delta\nu = 2\,\pi\alpha x = Su \text{ for } x < 0.2 \tag{6-12}$$

Thus, the absorptance for a single line increases linearly as the path length and the line intensity increases whenever the weak-line approximation is valid. Equation (6-12) is accurate within 10% when $x < 0.2$.

When either the path length is large or the pressure is small, the absorption may be complete over a wave-number region of several half-widths around the line center. In this case, the strong-line approximation is valid, which is equivalent to neglecting the factor unity compared to ν'^2 in the denominator of the exponential in Eq. (6-10). This can be done because the factor unity can always be neglected in the wings of the lines when $v' >> 1$. For other values of v', the exponential has a value very close to zero so that its exact value does not matter. When this factor is neglected [45-47],

$$A\,\Delta\nu = 2(S\alpha u)^{1/2} \text{ for } x > 1.63 \tag{6-13}$$

which is called the square root region, since the absorptance varies as the square root of the path length, pressure, and line intensity. Equation (6-13) is accurate within 10% when $x > 1.63$.

The integral in Eq. (6-10) can be evaluated exactly to obtain the following expression for the absorptance of a single spectral line [48]:

$$A\,\Delta\nu = 2\pi\alpha x e^{-x}\,[I_0(x) + I_1(x)] \tag{6-14}$$

where x is defined by Eq. (6-9) and I_0 and I_1 are the Bessel functions of imaginary argument. This function is tabulated in [45]. The limiting expressions given by Eq. (6-12) and (6-13) can be obtained from Eq. (6-14) from the usual expansions of the Bessel function for small and large values of the argument.

6.2.2. Single Line with Doppler Shape. When the spectral lines have the Doppler shape, the absorptance is obtained by substituting Eq. (6-4) into Eq. (6-1). When the weak-line approximation is valid, the absorptance is given by [49]

$$A\,\Delta\nu = Su \tag{6-15}$$

This is the same as Eq. (6-12); thus, the line shape does not affect the absorption when the lines are weak. Only the total line strength is of importance in this limiting case.

When the strong-line approximation is valid,

$$A\,\Delta\nu = 2\,\Delta\nu_D\left(\frac{\ln x_D}{\ln 2}\right)^{1/2} \qquad \text{for } x_D >> 1 \tag{6-16}$$

where

$$x_D = \left(\frac{\ln 2}{\pi}\right)^{1/2}\frac{Su}{\Delta\nu_D} \tag{6-17}$$

The absorptance increases very slowly as the number of absorbing molecules increases because the Doppler line shape drops off exponentially in the wings of the line. When the path is sufficiently long to absorb most of the radiation near the center of the line, it is not possible to absorb much additional radiation in the wings of the line by increasing the amount of absorbing gas. For long paths, the Doppler line shape acts qualitatively as though it absorbed all of the radiation over a bandwidth of $2\Delta\nu_D$ and none outside this interval. By contrast, the absorptance increases as the square root of the path length for the Lorentz line shape since the line shape varies as $(\nu - \nu_0)^{-2}$ in the wings.

When neither of these limiting expressions for the absorptance by a line with the Doppler shape is valid, it is necessary to use one of the more general expressions that have been derived or one of the many tables or graphs that have been calculated [43].

6.2.3. Single Line with Both Doppler and Lorentz Broadening. When the line shape is a result of both Doppler and Lorentz broadening (Eq. 6-6), the absorptance is still given by Eq. (6-15) when the weak-line approximation is valid. When the strong-line approximation is valid, the absorptance is given by [49]

$$A\,\Delta\nu = 2(S\alpha u)^{1/2}\left[\,1+\frac{3}{8}\left(1-\frac{2}{3}a^2\right)\frac{\pi\,\Delta\nu_D{}^2}{S\alpha u\,\ln 2}+\cdots\,\right] \tag{6-18}$$

where a is derived from Eq. (6-12). Higher-order terms are given in [49]. The leading term in this expression is the same as that of Eq. (6-13), which is the expression for the absorptance of a line with the Lorentz line shape in the square root region. In the strong-line region, the Lorentz line shape is much more important in determining the absorptance than the Doppler line shape because the Lorentz shape falls off much more slowly in the far wings of the line. The Doppler half-width may even be much larger than the Lorentz, yet the Lorentz line shape can still determine the absorptance.

More complicated expressions for the absorptance in intermediate regions are discussed in [43].

6.3. Absorption by Bands

When the spectral lines in a band do not overlap appreciably, the absorptance of a group of spectral lines can be calculated by summing the contribution from the individual spectral lines. However, in many cases of practical interest, the spectral lines overlap appreciably and this effect must be taken into account when the absorptance is calculated. When the lines overlap, the absorptance is always less than would be expected from the same number of isolated spectral lines.

The absorptance of a band of overlapping spectral lines depends on details of the relative spacing between the spectral lines and their intensity variation. Because of the many rapid variations of the absorption coefficient of a band as a function of frequency, it is very difficult to integrate Eq. (6-1), even with a large electronic computer. This is not necessary, however, since four models are available that represent the absorption from an actual band with reasonable accuracy. These four models are: (1) Elsasser model [50]; (2) statistical model [51, 52]; (3) random Elsasser model [46, 53]; (4) quasirandom model [54].

6.3.1. Elsasser Model. The Elsasser model assumes that the spectral lines are evenly spaced and that they all have the same intensity (Fig. 6-18*A*). Some portions of the CO_2 spectrum can be represented with fair accuracy by this model. However, there are always numerous weak lines between the stronger, regularly spaced lines in this spectrum and furthermore the line intensity varies with frequency. The weak lines absorb an increasing share of the radiation as the path length becomes longer. Thus the Elsasser model does not represent CO_2 absorption accurately over a wide range of path lengths.

The exact expression for the absorptance of an Elsasser band is [50]

$$A = 1-\frac{1}{2\pi}\int_{-\pi}^{\pi}\exp\frac{-\beta x\,\sinh\beta}{\cosh\beta-\cos z}\,dz \tag{6-19}$$

where

$$\beta = 2\pi\alpha/d \tag{6-20}$$

x is defined by (6-9), and d is the line spacing. This integral cannot be evaluated in closed form, but useful expressions for the absorptance can be found in certain limits (Sec. 6.4). An extensive calculation of the integral in Eq. (6-19), including a table

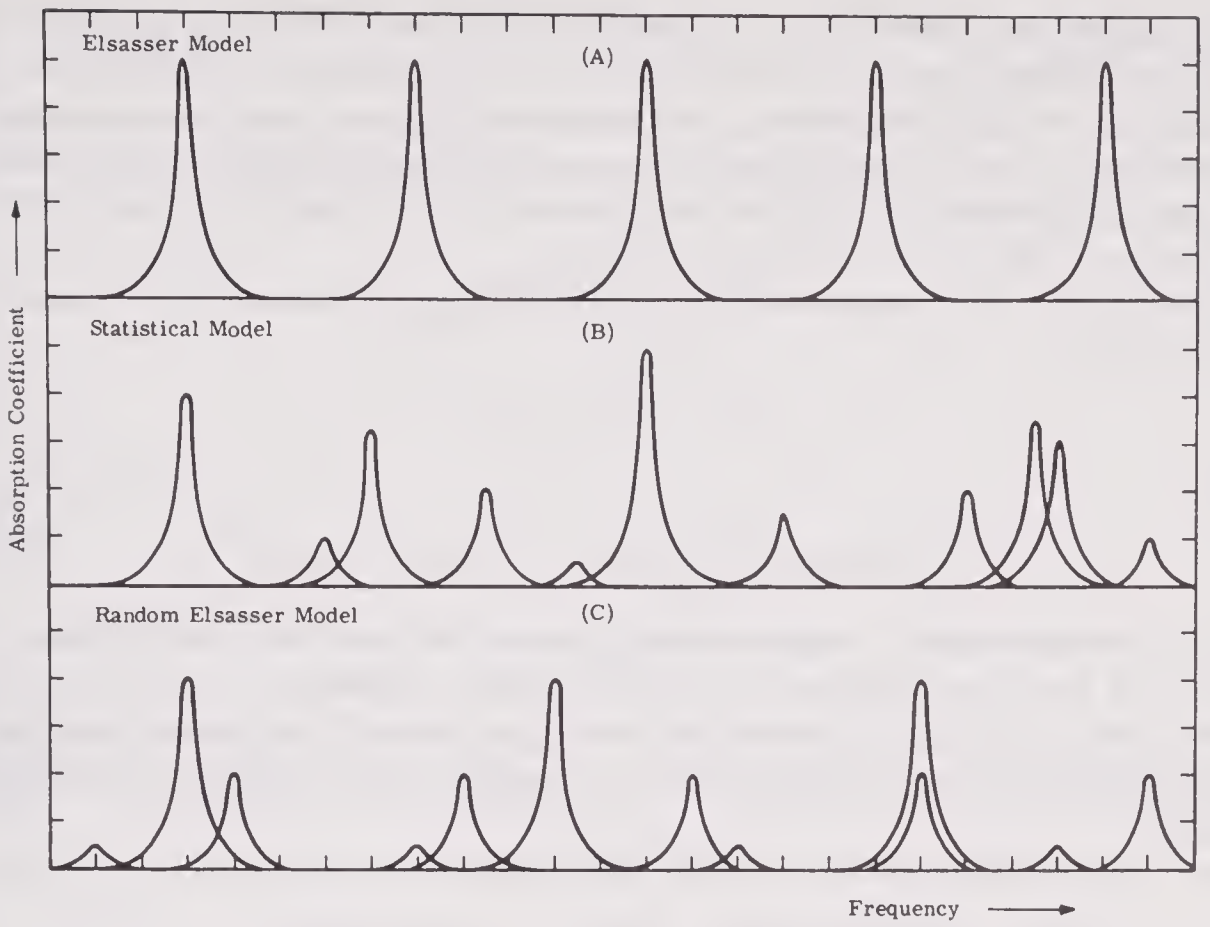


FIG. 6-18. Comparison of Elsasser, statistical and random Elsasser model.

for values of β from 0.0001 to 1.0 and for values of Su/d sinh β from 0.02 to $1.5(10)^5$ is given in [55]. The variation of the absorptance as a function of $\beta^2 x$ is shown in Fig. 6-19. The various curves are for particular values of β; *i.e.*, they give the absorptance at constant pressure. The linear and square root approximations are indicated to show their regions of validity.

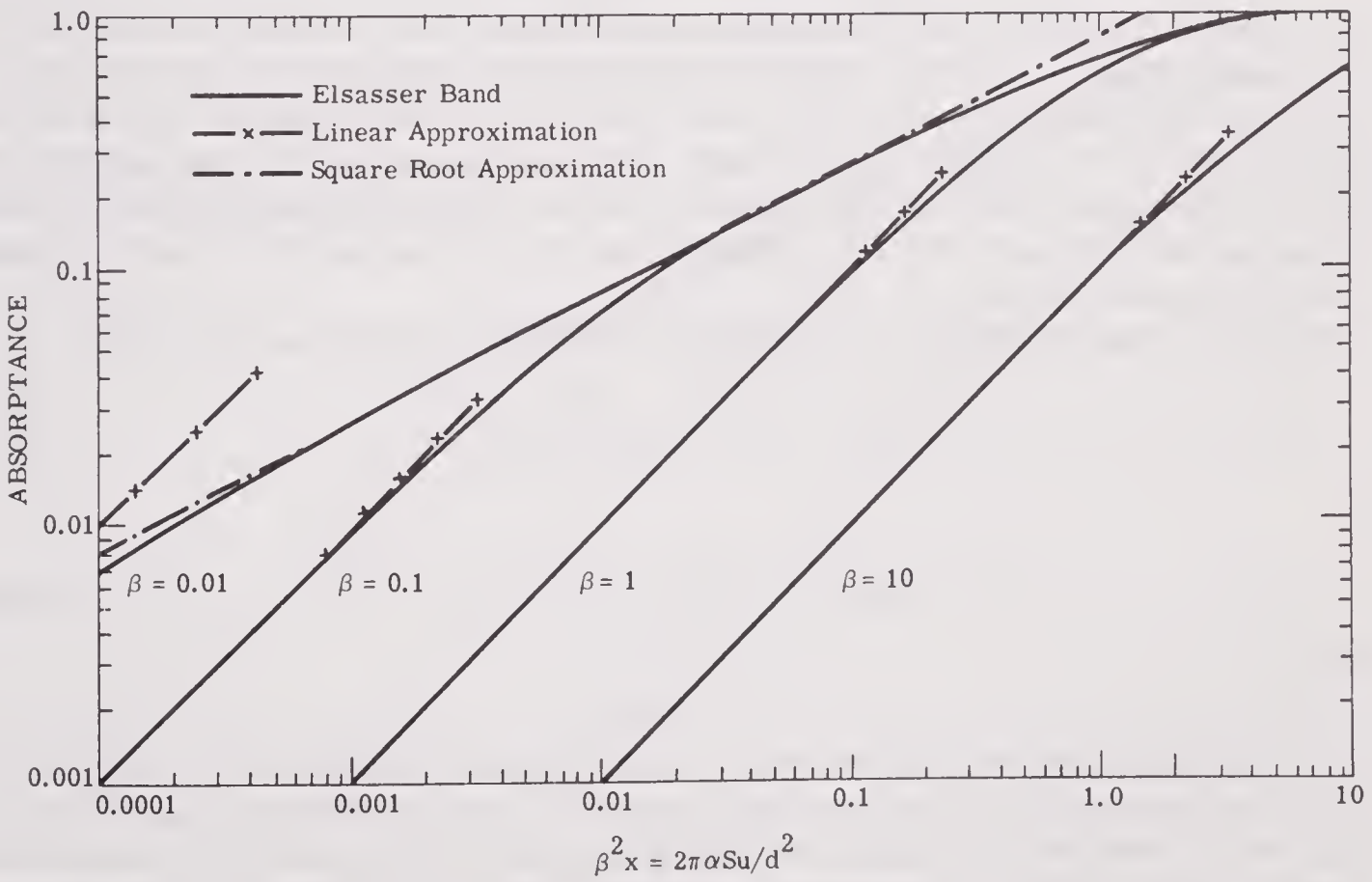


FIG. 6-19. Absorption for an Elsasser band as a function of $\beta^2 x = 2\pi\alpha Su/d^2$.

6.3.2. Statistical Model. The statistical or Mayer-Goody model assumes that the spectral lines have a random spacing (Fig. 6-18*B*) as contrasted to the regular spacing of the Elsasser model. The intensity of the spectral lines can vary in any manner whatsoever as long as it can be represented by some distribution function. The absorptance of H_2O can be represented by the statistical model over a moderate range of path lengths and pressures.

The absorptance over a wave-number interval, D, when the statistical model is valid, is given by [46]

$$A = 1 - (1 - \overline{A}_{sl,\,D})^n \tag{6-21}$$

where

$$\overline{A}_{sl,\,D}\,(S_0, \mu, p) = \int_0^\infty A_{sl,\,D}\,(S, \mu, p) P(S, S_0)\, dS$$

$A_{sl,D}$ is the absorptance of a single isolated spectral line over the wave-number interval D, $P(S,S_0)$ is the normalized probability of finding a spectral line with an intensity between S and $S + dS$, S_0 is a parametric mean line intensity that occurs in the intensity distribution function, and n is the number of spectral lines in the interval D whose mean spacing is d.

When n in the interval D is large, Eq. (6-21) approaches the form [46, 51, 52]

$$A = 1 - \exp\,(-n\overline{A}_{sl,\,D}) \quad \text{for } n >> 10 \tag{6-22}$$

Equation (6-22) should not be used when there are a small number of lines in the interval.

As examples of the use of these equations assume that: (*a*) all the spectral lines are equally intense, so that

$$P(S) = \delta(S - S_0) \tag{6-23}$$

(*b*) an exponential distribution of line intensities,

$$P(S) = S_0^{-1} \exp\,(-S/S_0) \tag{6-24}$$

For intensity distribution (*a*), the absorptance of a single line averaged over the distribution is obtained by replacing S by S_0 in the appropriate expression for the absorptance of a single line given (Sec. 6-2). This result can then be substituted into Eq. (6-21) or (6-22) to obtain the absorptance of a band. The absorptance for this case is shown in Fig. 6-20, assuming that all of the spectral lines are equally intense, that the Lorentz line shape is valid, and that the absorption is from a large number of spectral lines with an average spacing d.

The single-line absorptance as calculated for intensity distribution (*b*) is [46]

$$A = 1 - \left[\, 1 - \frac{\beta x_0}{n(1 + 2\,x_0)^{1/2}} \,\right]^n \tag{6-25}$$

or

$$A = 1 - \exp\left[\, -\frac{\beta x_0}{(1 + 2\,x_0)^{1/2}} \,\right] \quad \text{for } n >> 10 \tag{6-26}$$

where

$$x_0 = S_0 \mu / 2\pi\alpha$$

The expressions for the absorptance of a single line that are used in deriving Eq. (6-25) and (6-26) assume an infinite frequency interval. These approximate expressions for $A_{sl,D}$ may be used only so long as there is no appreciable absorption by the single line outside of the interval D. In general, more complicated expressions for the absorptance

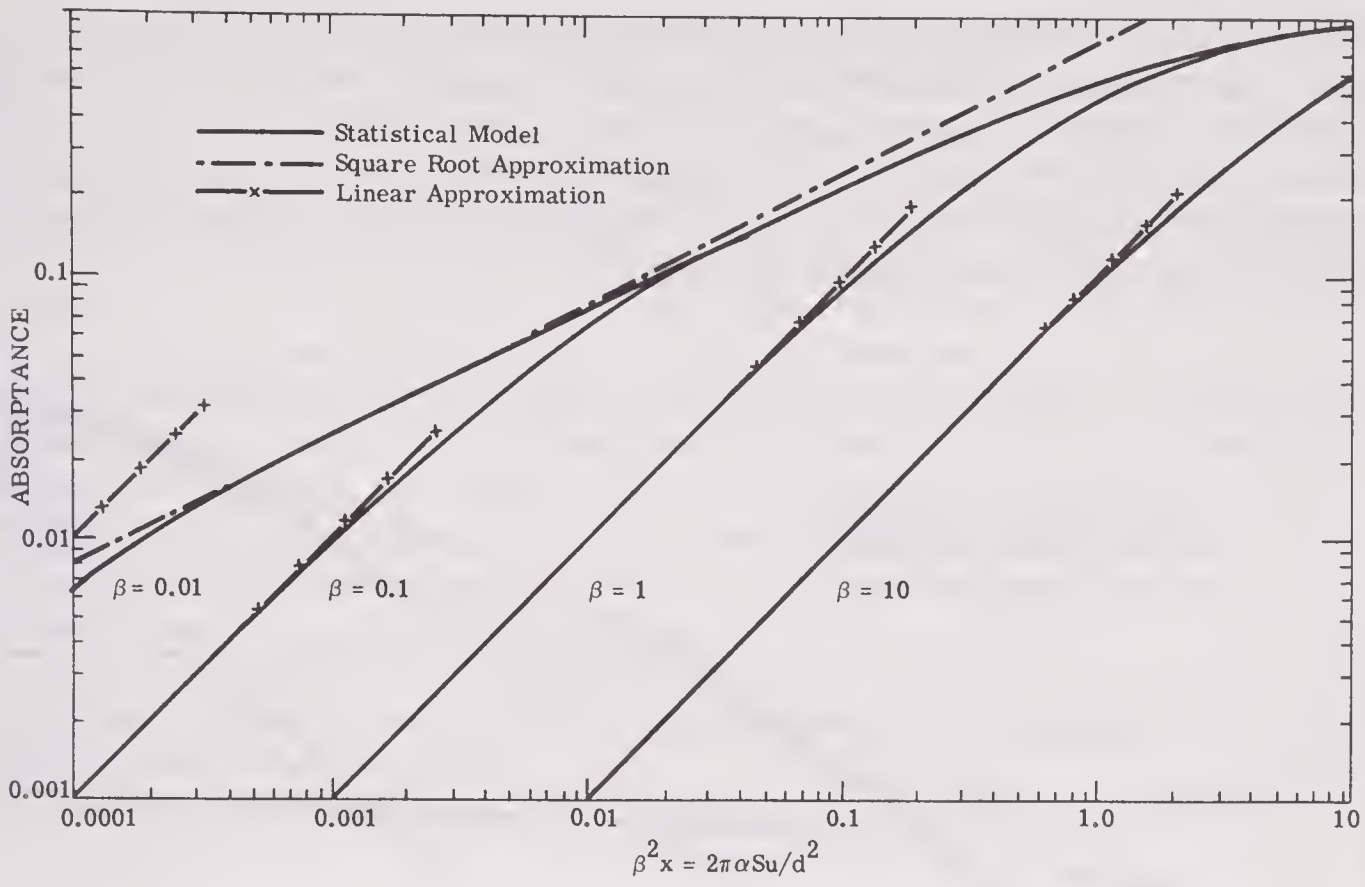


FIG. 6-20. Absorption for the statistical model as a function of $\beta^2 x$.

of a single line over a finite frequency interval must be used [54]. The single-line expressions for the absorptance of a line with the Doppler shape or the Doppler-Lorentz shape can also be substituted into Eq. (6-21) or (6-22).

The absorptance for the statistical and Elsasser models is compared in Fig. 6-21. The absorptance is always larger at long path lengths for the Elsasser model since the spectral lines absorb most efficiently with regular spacing. With a random spacing, the lines overlap more strongly because lines which happen to be close together cannot absorb as efficiently as when they are spaced further apart.

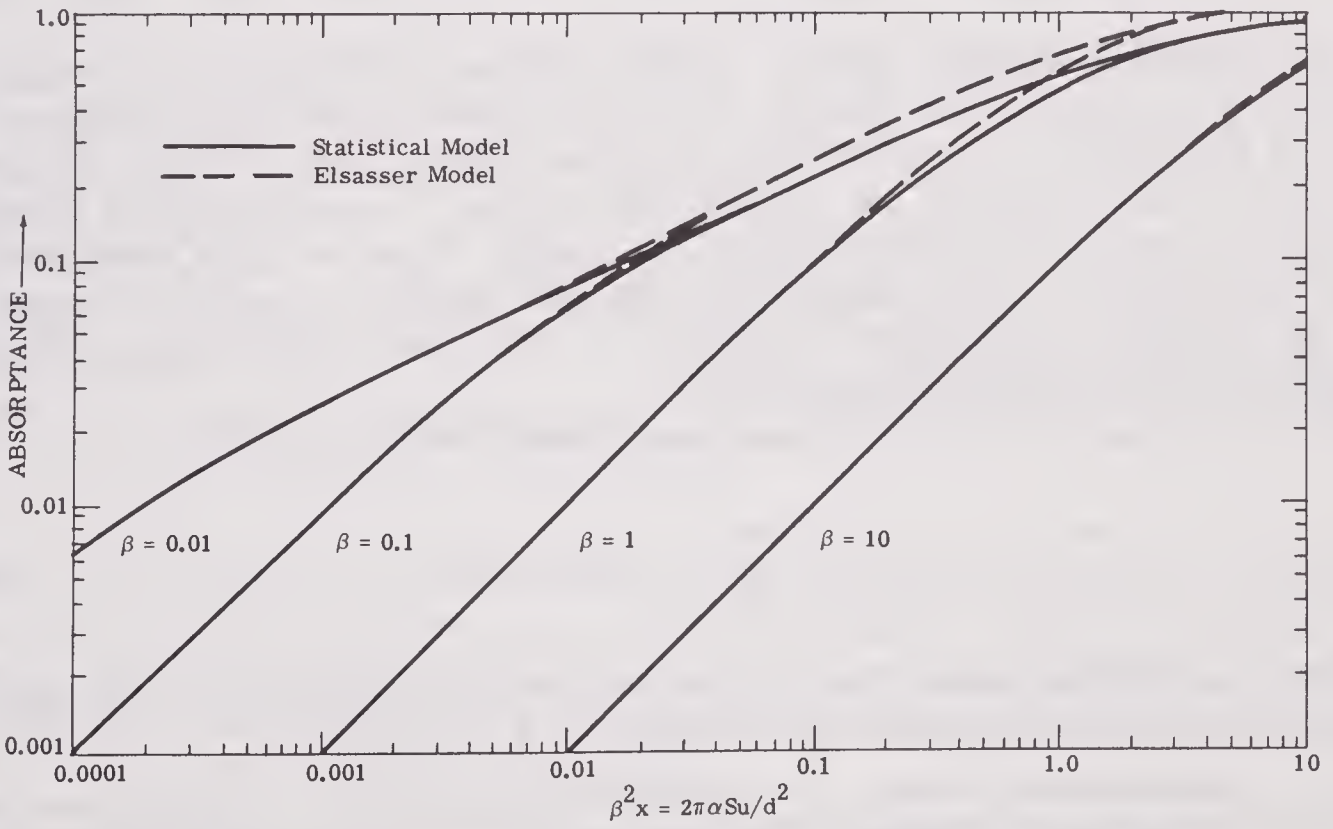


FIG. 6-21. Comparison of absorption from the statistical and Elsasser band models.

The absorptance for the Lorentz, Doppler, and square (absorption coefficient constant over a finite frequency interval) line shape are compared in Fig. 6-22 for a random distribution of lines. In each case, the statistical model with all of the lines equally intense is assumed. For comparison purposes, β_D and x_D are set equal to β and x for the Lorentz line shape. The very slow increase of the absorptance for large path lengths can be seen for the Doppler line shape.

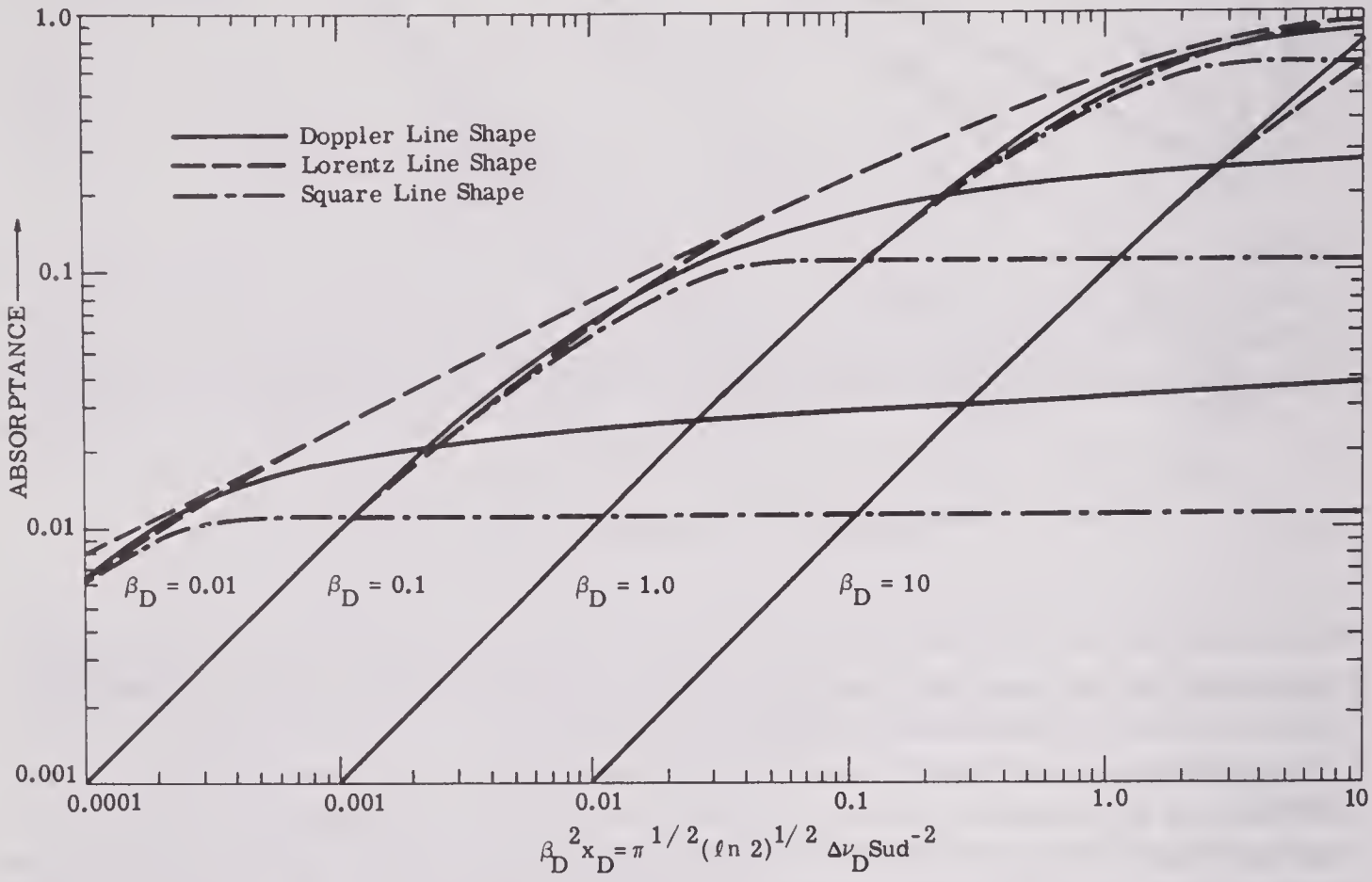


FIG. 6-22. Absorption of spectral lines with Doppler, Lorentz, and square line shape. Statistical model with equally intense lines assumed.

6.3.3. Random Elsasser Model. A more accurate representation of band absorption is provided in many cases by the random Elsasser model, which assumes the random superposition of several different Elsasser bands. Each of these superposed bands may have a different line intensity and spacing. As many different Elsasser bands as desired may be superimposed in this model. Thus, all of the weak spectral lines that contribute to the absorption for the path lengths and pressures considered can be included in the absorption calculations. The superposition of three different Elsasser bands is illustrated in Fig. 6-18*C*.

The absorptance of M randomly superposed Elsasser bands is [46, 53]

$$A = 1 - \prod_{i=1}^{M} [1 - A_{E,i}\ (x_i, \beta_i)] \tag{6-27}$$

where $A_{E,i}$ is the absorptance of an Elsasser band with a half-width of α_i, a line spacing d_i, and a line intensity S_i, so that $x_i = S_i u/2\pi\alpha_i$ and $\beta_i = 2\pi\alpha_i/d_i$.

6.3.4. Quasirandom Model. The quasirandom model is the most accurate and, necessarily, the most complicated of the band models. It is especially useful when the absorptance is required over a wide range of path lengths and pressures. The spectral

lines in an actual band are arranged neither as regularly as required by the Elsasser band nor in as random a fashion as in the statistical model; there is some order in their arrangement. In the quasirandom model the absorptance is calculated first for a frequency interval that is much smaller than the interval size of interest. This localizes the stronger lines to a narrow interval around their actual positions and prevents the introduction of spurious overlapping effects. The absorptance of this narrow interval is calculated from the equation for the single-line absorptance over a finite interval [54]. The absorptance for each of the N spectral lines in the interval is calculated separately and the results combined by assuming a random placing of the spectral lines within the small interval. The absorption from the wings of lines in neighboring intervals is included in the calculation. The results are averaged for at least two different arrangements of the mesh that divides the spectrum into frequency intervals. Finally the absorptance values for all of the small intervals that fill the larger interval of interest are averaged to obtain the final value for the absorptance. An electronic computer is commonly used to calculate results for this model when many spectral lines are involved. The many weak spectral lines and their relative spacing are accurately taken into account by this model.

The absorptance for the quasirandom model is given by [54]

$$A = \frac{1}{L} \sum_{j=1}^{L} A_j \tag{6-28}$$

where A_j is the absorptance of each of the L smaller wave-number intervals into which the original interval $\Delta\nu$ is subdivided. The absorptance A_j is calculated from

$$A_j = 1 - \prod_{i=1}^{M} \left[1 - A_{sl,D}{}^{(i)}\ (S_i, d_i)\right] \tag{6-29}$$

where $A_{sl,D}{}^{(i)}$ is the single-line absorptance over the finite interval D for a line with intensity S_i and half-width α_i, and M is the number of lines in the frequency interval j.

References [56] and [57] give absorptance tables for H_2O and CO_2 based on the quasirandom model calculations. In these calculations, the lines in the small frequency interval were grouped by intensity decades. The average intensity and number of lines in each of these decades were calculated and used in Eq. (6-29). All lines from all isotopic species having an intensity greater than 10^{-8} of the strongest line in an absorption region were included in the calculation. The final results for the transmissivity are given every 5 cm^{-1} for CO_2 from 500 to 10,000 cm^{-1} and for H_2O from 1000 to 10,000 cm^{-1}. Values averaged over 20, 50, and 100 cm^{-1} intervals are also presented. All values are given at three temperatures, 200°, 250°, and 300°K, and for seven pressures, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, and 1 atm. The path length for CO_2 ranges from 0.2 to 10,000 atm cm and for H_2O from 0.001 to 50 cm of precipitable water.

6.4. Useful Approximations to Band Models

There are three different limits in which approximate forms of the equations given in Sec. 6.2 and 6.3 can be obtained. These are: (1) weak-line approximation; (2) strong-line approximation; (3) nonoverlapping-line approximation. Each of the approximations is valid over a wide range of path lengths and pressures. However, care must be taken to use each approximation only within its own region of validity. A detailed study of these approximations is given in [47].

6.4.1. Weak-Line Approximation. The weak-line approximation assumes that the absorption due to each line considered individually is small even at the line center.

However, the total absorptance due to a number of such lines may have any value (even near unity), since the sum of a number of small quantities is not necessarily small.

For the Elsasser band, Eq. (6-19) reduces to [47]

$$A = 1 - e^{-\beta x} \tag{6-30}$$

in the weak-line limit. When $\beta x << 1$, Eq. (6-30) reduces to Eq. (6-12). When this inequality is satisfied, the lines do not overlap and the absorptance increases linearly with path length. When the lines begin to overlap, the exponential term in Eq. (6-13) takes account of this effect. On a logarithmic plot of absorptance (Fig. 6-19), the slope of the absorptance curve is unity in the linear region where the lines do not overlap and the weak-line approximation is valid. When this approximation continues to be valid as the path length increases, the slope of the absorptance curve decreases, as shown on the curves marked $\beta = 1$ or 10, as the lines begin to overlap.

The regions of validity of this approximation are shown in Fig. 6-23 and Table 6-4. Within the indicated regions, the approximations are accurate within 10%.

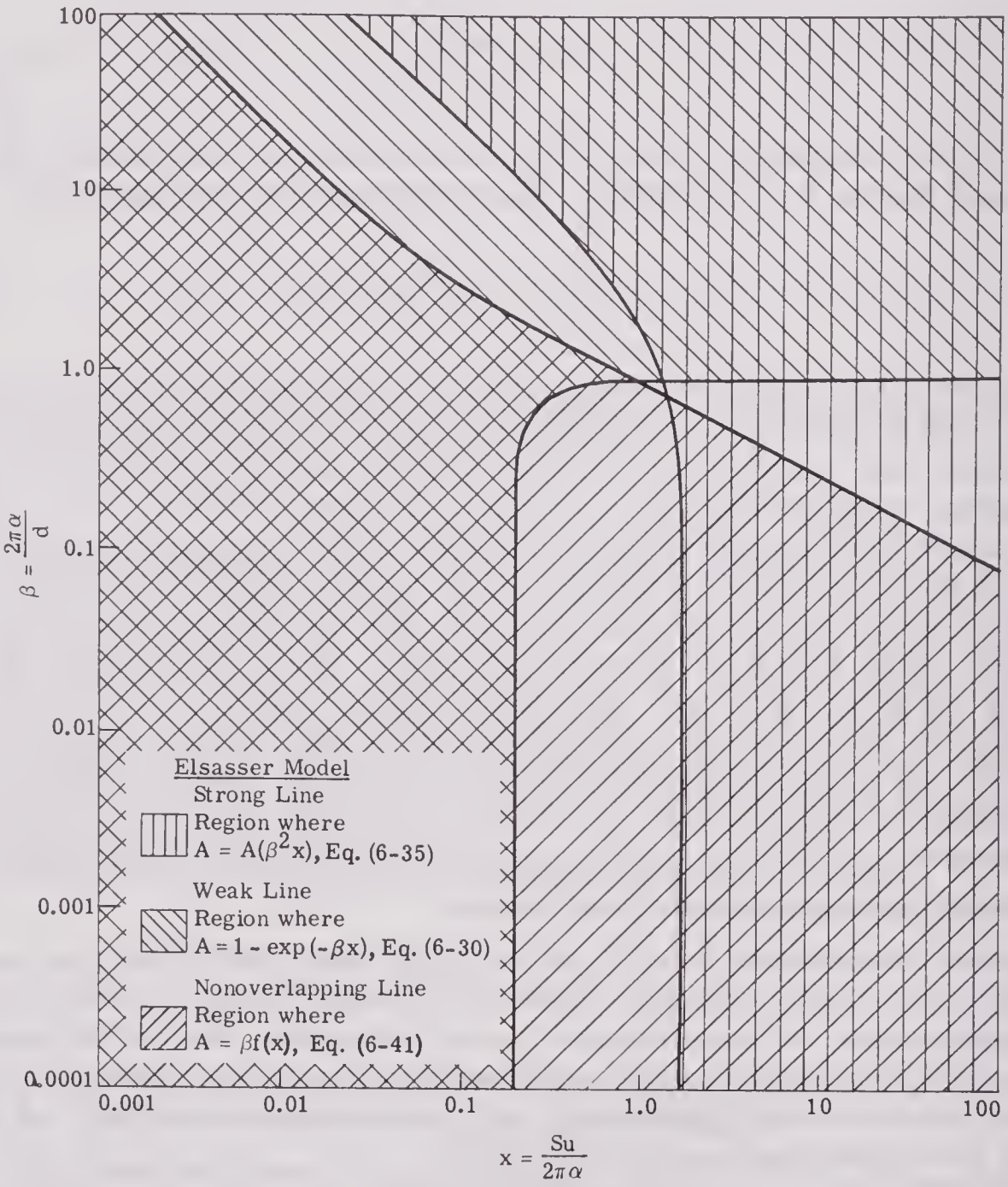


FIG. 6-23. Regions of validity for band model approximations (Elsasser model).

TABLE 6-4. REGIONS OF VALIDITY OF VARIOUS APPROXIMATIONS FOR BAND ABSORPTION.[a]

Approximation	$\beta = 2\pi\alpha/d$	*Elsasser Model*	*Statistical Model; All Lines Equally Intense*	*Statistical Model; Exponential Line Intensity Distribution*
Strong-line approximation	0.001	$x > 1.63$	$x > 1.63$	$x_0 > 2.4$
	0.01	$x > 1.63$	$x > 1.63$	$x_0 > 2.4$
	0.1	$x > 1.63$	$x > 1.63$	$x_0 > 2.3$
	1	$x > 1.35$	$x > 1.1$	$x_0 > 1.4$
	10	$x > 0.24$	$x > 0.24$	$x_0 > 0.27$
	100	$x > 0.024$	$x > 0.024$	$x_0 > 0.24$
Weak-line approximation	0.001	$x < 0.20$	$x < 0.20$	$x_0 < 0.10$
	0.01	$x < 0.20$	$x < 0.20$	$x_0 < 0.10$
	0.1	$x < 0.20$	$x < 0.20$	$x_0 < 0.10$
	1	$x < \infty$	$x < 0.23$	$x_0 < 0.11$
	10	$x < \infty$	$x < \infty$	$x_0 < \infty$
	100	$x < \infty$	$x < \infty$	$x_0 < \infty$
Nonoverlapping line approximation	0.001	$x < 600\,000$	$x < 63\,000$	$x_0 < 80\,000$
	0.01	$x < 6000$	$x < 630$	$x_0 < 800$
	0.1	$x < 60$	$x < 6.3$	$x_0 < 8$
	1	$x < 0.7$	$x < 0.22$	$x_0 < 0.23$
	10	$x < 0.02$	$x < 0.020$	$x_0 < 0.020$
	100	$x < 0.002$	$x < 0.0020$	$x_0 < 0.0020$

[a]When $x = Su/2\pi\alpha$ satisfies the given inequalities, the indicated approximation for the absorption is valid with an error of less than 10%. For the exponential line-intensity distribution. $x_0 = S_0 u/2\pi\alpha$, where $P(S) = S_0^{-1} \times \exp(-S/S_0)$.

The weak-line approximation for the statistical model is [47]

$$A = 1 - \left(1 - \frac{\beta x}{n}\right)^n \tag{6-31}$$

or

$$A = 1 - e^{-\beta x} \tag{6-32}$$

These equations follow for any intensity distribution. For example, when $P(S)$ is given by Eq. (6-24), replace x by x_0 in equations 6-31 and 6-32. The regions of validity of this approximation are given in Table 6-4 and are shown in Fig. 6-24. Within the indicated regions, the approximations are accurate within 10%.

The weak-line approximation for the random Elsasser model is [47]

$$A = 1 - \prod_{i=1}^{M} e^{-\beta_i x_i} \tag{6-33}$$

and for the quasirandom model is [54]

$$A = \frac{1}{L} \sum_{j=1}^{L} \left[1 - \prod_{i=1}^{M} (1 - e^{-\beta_i x_i}) \right] \tag{6-34}$$

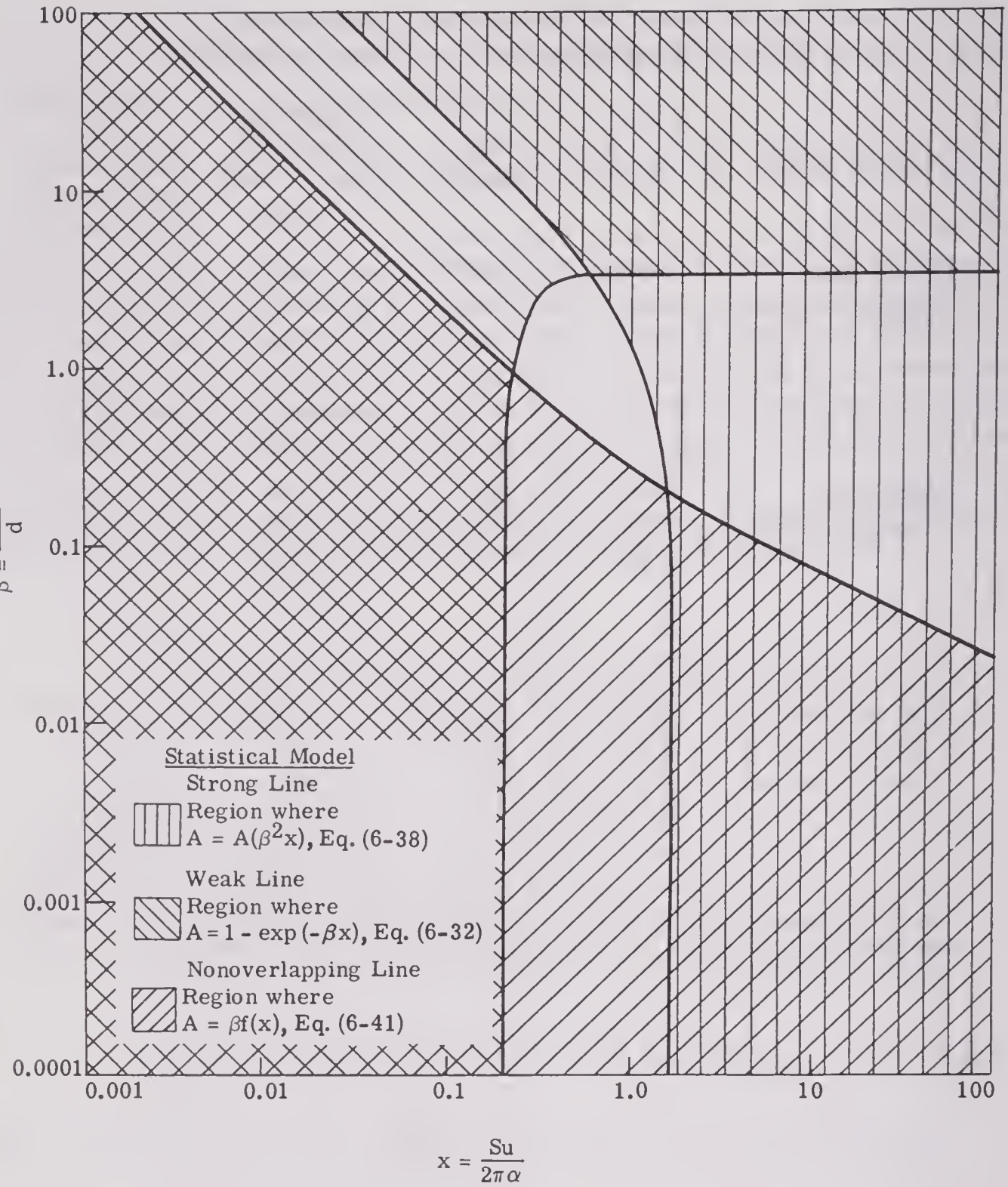


FIG. 6-24. Regions of validity for band model approximations (statistical model).

6.4.2. Strong-Line Approximation. The strong-line approximation assumes that the absorption is virtually complete over a bandwidth of at least several half-widths around the line center.

For the Elsasser model, the strong-line approximation is [45]

$$A = \phi\left[\left(\frac{1}{2}\beta^2 x\right)^{1/2}\right] \tag{6-35}$$

where ϕ is the error function integral defined as

$$\phi(z) = \frac{2}{\pi^{1/2}}\int_0^z e^{-z^2}\,dz \tag{6-36}$$

When $\beta^2 x << 1$, Eq. (6-35) reduces to Eq. 6-13, which is the square root approximation (Fig. 6-19). For example, the curve marked $\beta = 0.01$ has a slope of nearly one-half over most of the region shown here. Beyond about $\beta^2 x = 0.05$, however, the absorptance curve begins to depart from the square root approximation as the spectral lines begin to overlap.

The rather severe limitations of Eq. (6-35) should always be considered. The band under consideration must have regularly spaced and equally intense lines. Furthermore, the equation is only valid in the strong-line region, as indicated in Fig. 6-21 and Table 6-4.

The strong-line approximation for the absorptance for the statistical model is [47]

$$A = 1 - \left[1 - \left(\frac{2\beta^2 x}{\pi}\right)^{1/2}\right]^N \tag{6-37}$$

or

$$A = 1 - \exp - (2\beta^2 x/\pi)^{1/2} \quad \text{for } N >> 10 \tag{6-38}$$

where an appropriately defined average value of the line intensity S is used in the determination of the value of x. Otherwise, the term in the square bracket of Eq. (6-37) may be evaluated for each of the N spectral lines and the results multiplied together. The regions of validity of this approximation are given in Fig. 6-24 and Table 6-4.

For the random Elsasser model, the strong-line approximation is [47]

$$A = 1 - \prod_{i=1}^{M} \left\{ 1 - \phi \left[\left(\frac{1}{2} \beta_i{}^2 x_i \right)^{1/2} \right] \right\} \tag{6-39}$$

where x_i and β_i are as defined after Eq. (6-27).

For the quasirandom model, the strong-line approximation is [46, 54]

$$A = \frac{1}{L} \sum_{j=1}^{L} \left[1 - \prod_{i=1}^{M} \left\{ e^{-z_i{}^2} - \pi^{1/2} z_i [1 - \phi(z_i)] \right\} \right] \tag{6-40}$$

where

$$z_i{}^2 = 8\alpha_i{}^2 x_i / D^2$$

The expression for the absorptance over a finite frequency interval D [46] has been used in deriving Eq. (6-40).

6.4.3. Nonoverlapping-Line Approximation. The regions of validity for the strong- and weak-line approximations are determined by whether the absorption is large or small at the frequency of the line centers and do not depend on the degree of overlapping of the spectral lines. On the other hand, the only requirement for the validity of the nonoverlapping-line approximation is that the spectral lines do not overlap appreciably. It is valid regardless of the value of the absorption at the line centers. Since there is no effect from the overlapping of the spectral lines, the absorption is independent of whether the spacing between the lines is regular, random, or quasirandom. However, the absorptance can depend on the distribution of line intensities within the band. The nonoverlapping line approximation is particularly useful for extrapolating the absorption to small values of path length and of the pressure.

For any model, the absorptance when the nonoverlapping approximation is valid is [47]

$$A\,\Delta\nu = \sum_{i=1}^{N} 2\pi\alpha_i x_i e^{-x_i}[I_0(x_i) + I_1(x_i)] \tag{6-41}$$

where the summation is over the N spectral lines in the frequency interval $\Delta\nu$. The regions of validity of this approximation are shown in Fig. 6-23 and 6-24 and are given in Table 6-4.

A particularly simple result is obtained if the probability distribution of line intensities is given by Eq. (6-24). Then the absorptance is [47]

$$A\,\Delta\nu = 2\pi\alpha x_0(1 + 2x_0)^{-1/2} \tag{6-42}$$

where

$$x_0 = S_0\mu/2\pi\alpha$$

6.5. Scattering

Pure scattering occurs if there is no absorption of the radiation in the process, and, hence, no loss of energy but only a redistribution of it. Most of the scattering encountered in the atmosphere is essentially pure. Typical exceptions to this are the dense black smoke issuing from a boiler in which there is incomplete combustion or from a volcano emitting large amounts of fly ash. In these cases, the carbon and mineral particles not only scatter but also absorb strongly. Only pure scattering is discussed in this section. Concepts developed for the visible region of the spectrum are used in extending the study of scattering into the infrared region.

6.5.1. Relationship to Field of View. The attenuation due to scattering of a collimated beam of light depends upon the field of view of the receiving instrument. If the field of view is very large, some light scattered at a very small forward angle will still be accepted and recorded; if the field of view is very small, virtually all scattered radiation can be rejected and only transmitted radiation registered.

The influence of the field of view of a measuring radiometer or telephotometer is discussed theoretically in [58]. Experimental investigations are described in [59] and [60]. The dependence of measured results on the field of view appears to be adequately described by the empirical relationship [60]

$$T_\theta = T + 0.5\ (1\text{-}T)\,(1\text{-}e^{-\theta})$$

where θ = angular diameter of radiometer field of view (radians)
T_θ = transmittance measured at a particular λ with a radiometer having a field of view θ
T = transmittance of unscattered light

This relationship agrees well with measurements when the meteorological range (Sec. 6.5.2) is greater than 10 km but deviates from measurements in a hazier atmosphere. Reference [61] presents tables based on the angular scattering function of a normal atmosphere [62], from which the scattered component can be evaluated for a given state of the atmosphere. A comparison of predicted values with measured values is shown in Fig. 6-25.

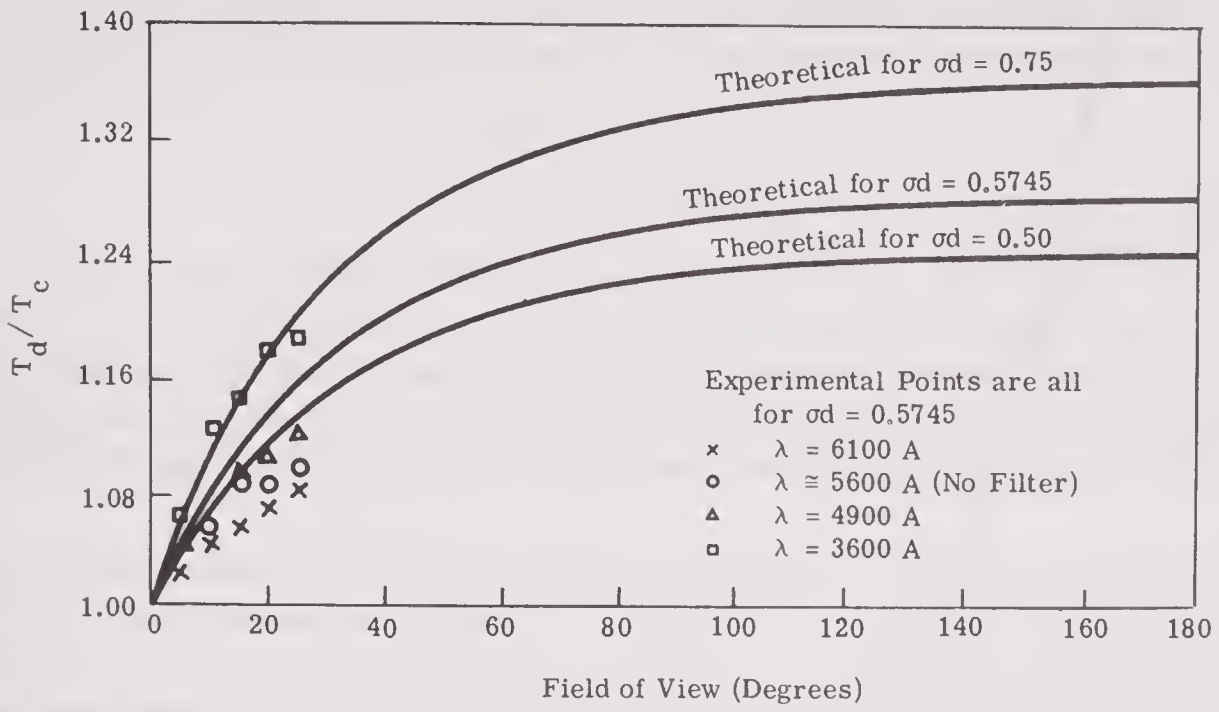


FIG. 6-25. Diffuse-to-collimated transmission ratio, T_d/T_c, as a function of field of view (σd = optical density of transmission path) [56].

6.5.2. Meteorological Range. Meterological range is a convenient parameter for visually describing a particular scattering condition. It is defined as the distance at which the average eye can just barely detect a large, black (nonreflecting and nonemissive) target against the horizon sky. To obtain a definition independent of individual eyes, meteorological range is further defined as a 2% contrast between the distant black target and the sky. The 2% figure represents a good average for the human eye's capabilities. The meteorological range, on the basis of the 2% contrast figure, is related to the visible scattering (Sec. 6.5.3) coefficient, σ (km^{-1}) by the relationship:

$$V = \left(\frac{1.0}{0.02}\right)\left(\frac{1}{\sigma n}\right) = 3.9/\sigma \tag{6-43}$$

where V = meteorological range (km)
σ = scattering coefficient (km^{-1})

The scattering coefficient is related to the transmission (or transmissivity) of a given optical path by the relationship:

$$T = e^{-\sigma x} \tag{6-44}$$

where T = transmission of optical path of length, x (dimensionless)
x = optical path length (km)
σ = scattering coefficient (km^{-1})

Figure 6-26 shows meteorological range *versus* scattering coefficient. Consider a day on which the meteorological range is 20 km. Refer to Fig. 6-26; the visible scattering coefficient for this day is 0.0195/km. The transmission over a 5-km path is then:

$$T = e^{-0.195\times5} = 0.377 = 37.7\%$$

A detailed discussion and methods of estimating meteorological range are given in Chapters 6 through 9 of [58].

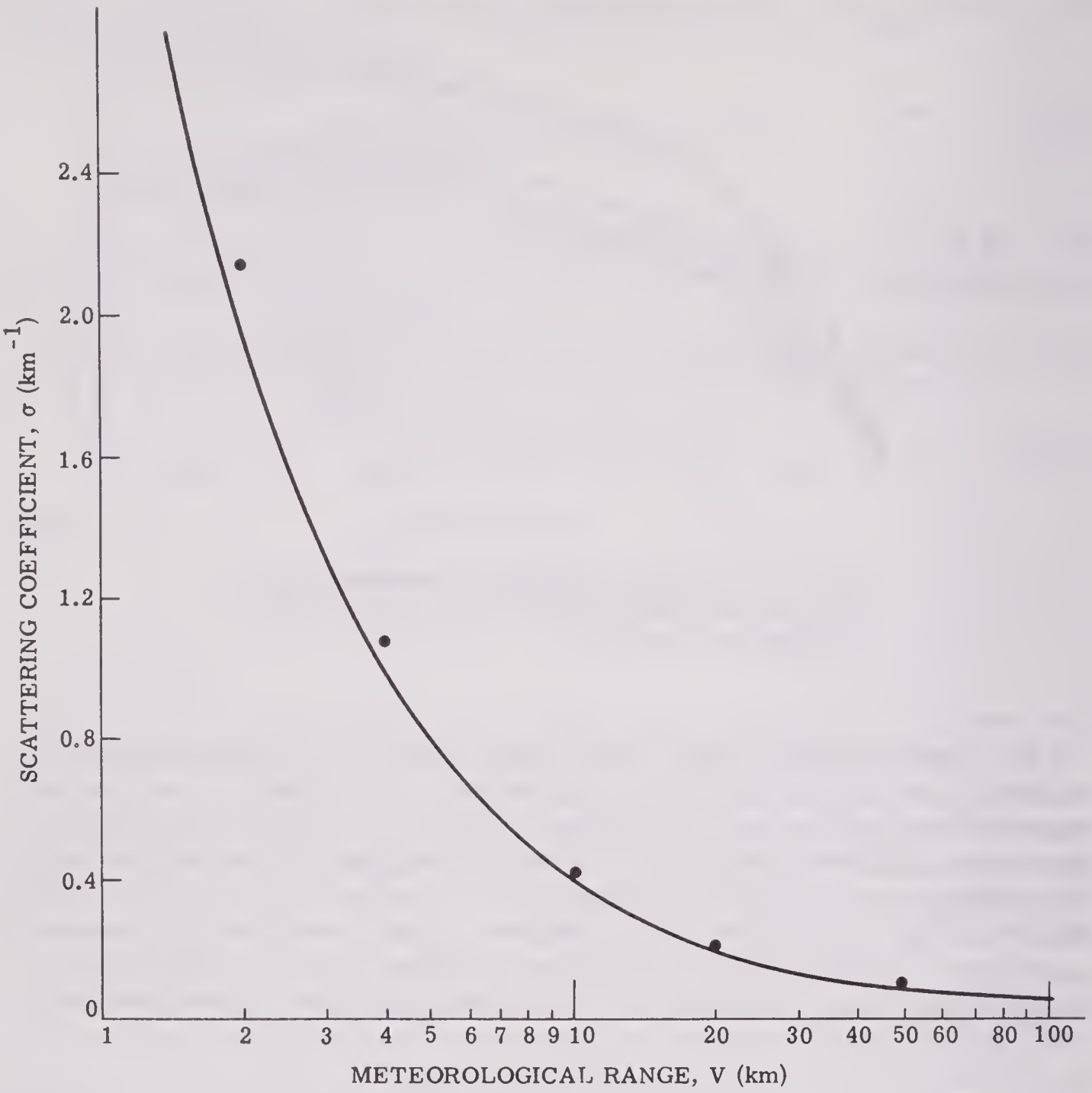


FIG. 6-26. Visible scattering coefficient, G, as a function of meteorological range. Solid points are visual ranges taken from the international visibility code.

6.5.3. Scattering Coefficient. Scattering can be treated theoretically in three separate approaches according to the relationship between the wavelength of the radiation being scattered and the size of the particles causing the scatter. These approaches are: (1) Rayleigh scattering; (2) Mie scattering, and (3) nonselective scattering.

6.5.3.1. *Rayleigh Scattering.* Rayleigh scattering applies when the radiation wavelength is much larger than the particle size. The volume scattering coefficient for Rayleigh scattering can be expressed as [63]

$$\sigma = (4\pi^2 NV^2/\lambda^4)\,(n^2 - n_0^2)^2/(n^2 + 2n_0^2)^2 \tag{6-45}$$

where N = number of particles per unit volume (cm^{-3})
V = volume of scattering particle (cm^3)
λ = wavelength of radiation (cm)
n_0 = refractive index of medium in which particles are suspended
n = refractive index of scattering particles

For spherical water droplets in air ($n_0 \cong 1$; $n \cong 1.33$ for the visible and near infrared, except in the vicinity of absorption bands where anomalous dispersion is encountered), Eq. (6-45) becomes

$$\sigma = 0.827NA^3/\lambda^4 \tag{6-46}$$

where A = the cross-sectional area of the scattering droplet. This expression must be integrated over the range of λ and A encountered in any given circumstance. As long as the original requirement is met for all λ and A; *i.e.*, the particle diameter ($2\sqrt{A/\pi}$) is very small compared to λ, the same scattering can be experienced from a large number of small particles or a small number of large particles, provided the product NA^3 is the same.

6.5.3.2. *Mie Scattering.* Mie scattering is applicable where the particle size is comparable to the radiation wavelength. The Mie scattering area coefficient is defined as the ratio of the area of the incident wave front that is affected by the particle to the cross-sectional area of the particle itself. The form of relationship between scattering-area coefficient and particle-size parameter is shown in Fig. 6-27. The value of K rises from 0 to nearly 4 and asymptotically approaches the value 2 for large droplets. The scattering coefficient, σ, is related to K by

$$\sigma = NK\pi a^2$$

or, for the almost universal condition in which there is a continuous size distribution in the particles, by

$$\sigma_\lambda = \pi \int_{a_1}^{a_2} N(a)K(a, n)a^2\, da \tag{6-47}$$

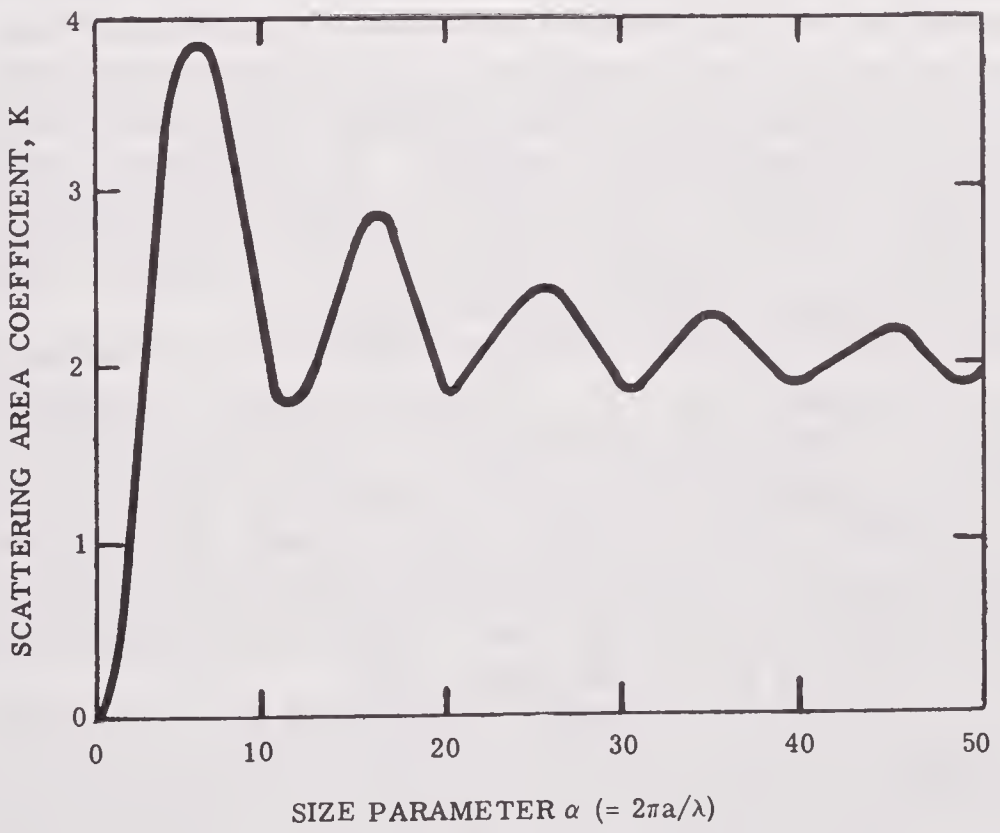


FIG. 6-27. Scattering area coefficient *vs.* size parameter for spherical water droplets (radius = a). Index of refraction, $n = 1.33$ [67].

where σ_λ = scattering coefficient for wavelength
$N(a)$ = number of particles per cubic centimeter in the interval da
$K(a, n)$ = scattering area coefficient
a = radius of spherical particle
n = index of refraction of particle

With the units of a and N in cm and cm^{-3}, respectively, σ is in units of cm^{-1}. To convert to km^{-1}, the more commonly used units for σ, multiply by 10^5.

References [64] and [65] present a detailed treatment of scattering theory. Reference [66] discusses application of theory for a wide variety of particle composition, size, and shape. Tables of the Mie scattering area coefficient are given in [67-69].

6.5.3.3. *Nonselective Scattering.* Nonselective scattering occurs when the particle size is very much larger than the radiation wavelength.

Large-particle scattering is composed of contributions from three processes involved in the interaction of the electromagnetic radiation with the scattering particle: (1) reflection from the surface of the particle with no penetration; (2) passage through the particle with and without internal reflections; and (3) diffraction at the edge of the particle. References [70], [71], and [72] discuss the combined effect of all three processes, including the interference encountered between the three components, and show that, for particles larger than about 2 times the wavelength of the radiation ($\alpha > 20$), the scattering-area coefficient becomes 2, which is the asymptotic value approached by the Mie coefficient. Thus, the theoretical approach through diffraction, refraction, and reflection appears to have little contribution to the more general approach of Mie. For $\alpha < 20$, the Mie theory is valid, and for $\alpha > 20$ the two predictions converge on the value 2. A generalized treatment of scattering considered as the sum of the refracted, diffracted, and reflected components is given in [58] and [73].

6.5.4. Scattering Coefficient Measurements. Reliable scattering-coefficient data in the infrared are difficult to obtain because of the contributions of both scattering and selective absorption to the extinction coefficient, which is the measured experimental value. The procedure usually followed is to confine the measurements to wavelengths that are as free as possible from absorption, *i.e.,* in the clearest part of the atmospheric windows between the major infrared absorption bands, and assume that the extinction coefficient measured is the scattering coefficient. As wavelength increases, however, it becomes virtually impossible to find regions completely free of absorption. A further assumption, which is valid for water droplets, is that absorption by the scattering droplets in the scattering process is negligibly small.

Figures 6-28 through 6-30 show scattering coefficient as a function of wavelength measured over Chesapeake Bay [74]. These curves show a persistent tendency to flatten out past 2 μ, which cannot be explained on the basis of any hypothetical particle-size distribution. It appears that, past 2 μ, an appreciable amount of absorption by the wings of the atmospheric absorption bands is present along with the scattering. There is a very slight tendency for the curves obtained on days of high absolute humidity to flatten out more rapidly than those obtained on days of low humidity, which indicates that the principal absorber is probably water vapor.

Figure 6-31 is the scattering coefficient measured over a 27.7-km path, essentially horizontal, at 10,000-ft elevation between two mountain peaks on the Island of Hawaii [74]. The atmosphere was usually very clear, and meteorological ranges of about 200 mi were common. The tendency to flatten out past 2 μ is again evident. For comparison, the values measured on the Nevada desert, another clear atmosphere, are included [75].

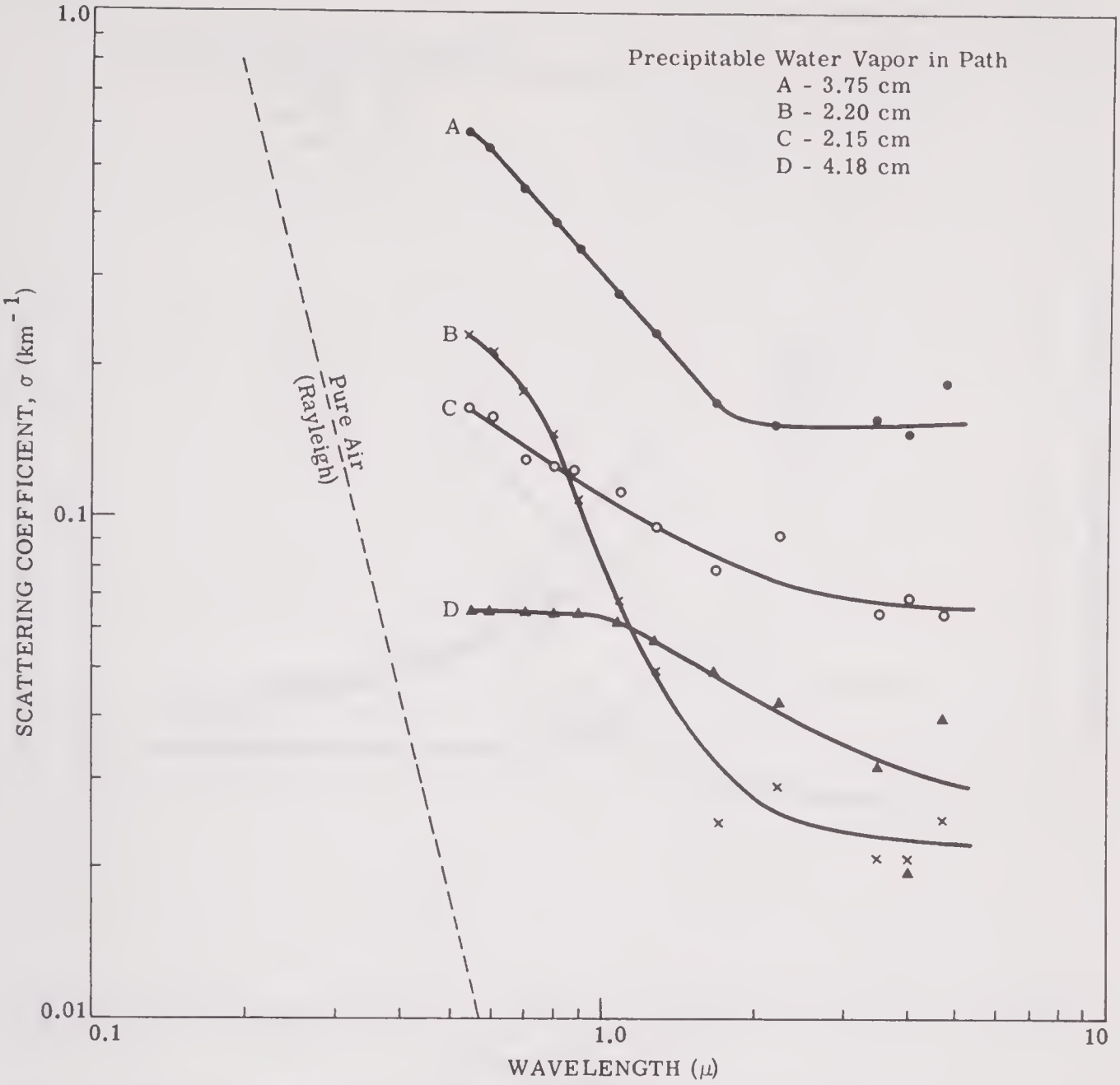


FIG. 6-28. Scattering coefficient measured over a 5.5-km sea-level maritime path [74].

Figures 6-32 and 6-33 show the results of additional scattering measurements over Chesapeake Bay [76]. Figure 6-32 contains curves of eight separate measurements made on days when the meteorological range extended from 100 km down to 3 km. The spread in points for the clearer days is attributable to severe scintillation (twinkle), which is normally encountered in a very clear atmosphere. Twinkle can also occur in a moist atmosphere but is usually masked. Figure 6-33 gives data for various meteorological ranges.

Figure 6-34 gives the scattering coefficients of a haze, and Fig. 6-35 gives the scattering coefficient and optical density of several types of fogs [38]. Optical density, D, and scattering coefficient, σ, are related by $\sigma = 2.3D$. The data are based on more than 600 spectrophotometric curves of hazes (optical density per km < 2), small drop fogs (optical density per km < 10), evolving fogs (changing size-distribution characteristics), and selective fogs and artificial smokes. The data indicate that stable fogs, with a density/km between 2 and 20, exhibit essentially the same opacity at all wavelengths out to 10 μ, at which point they invariably become sensibly more transparent. Evolving fogs and selective fogs, in which the visible density/km again could range from 2 to 20, showed a distinctly declining density as the wavelength increased.

The variation of scattering coefficient (visible) with height up to 10 km is shown in Fig. 6-36.

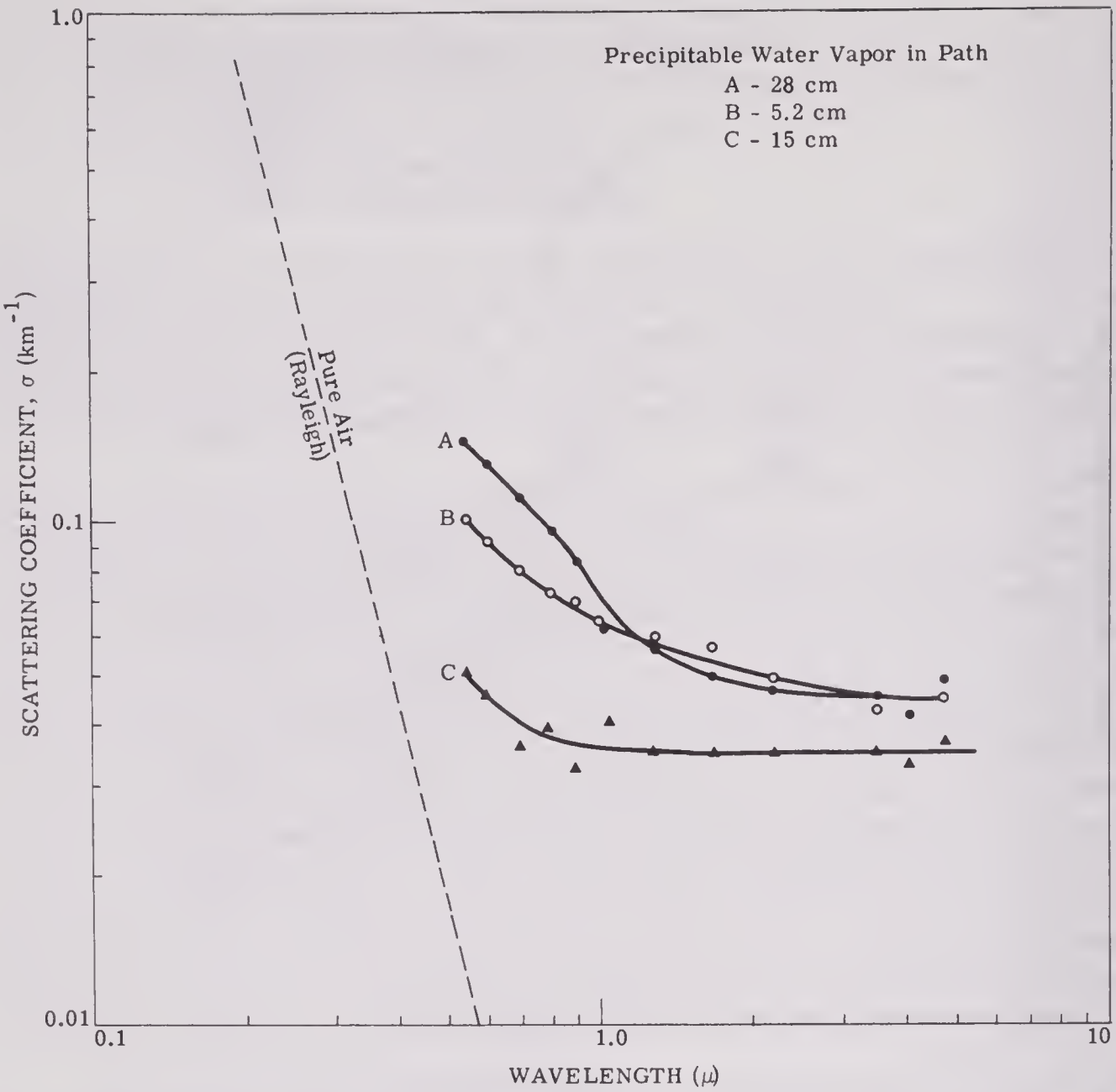


FIG. 6-29. Scattering coefficient measured over a 16.25-km sea-level maritime path. Curve *B* represents a typical winter day condition [74].

Because the scattering properties of the atmosphere can vary appreciably, it is not possible to state a scattering coefficient that will permit accurate predictions over a wide variety of conditions. However, a relationship frequently used is:

$$\sigma = c\lambda^{-\gamma} \tag{6-48}$$

where c and γ are constants determined by the concentration and size-distribution values for the aerosol, and λ is the wavelength of the radiations. If $\gamma = 4$, this relationship is recognized as the form of the Rayleigh coefficient. Actually, from the Mie theory, $\gamma = 4$ for very small particles, $\gamma = 2$ for a particle diameter equal to λ; $\gamma = 1$ for a particle diameter equal to $(3/2)\lambda$, and $\gamma = 0$ when the diameter is equal to 2λ.

Considering the presence of the atmospheric gases in the aerosol, a more exact form of this relationship is [77]

$$\sigma = c_1\lambda^{-\gamma} + c_2\lambda^{-4} \tag{6-49}$$

where the second term accounts for the scattering by the Rayleigh components. In most cases, however, the second term is considerably less than the first, and it may be neglected for practical purposes.

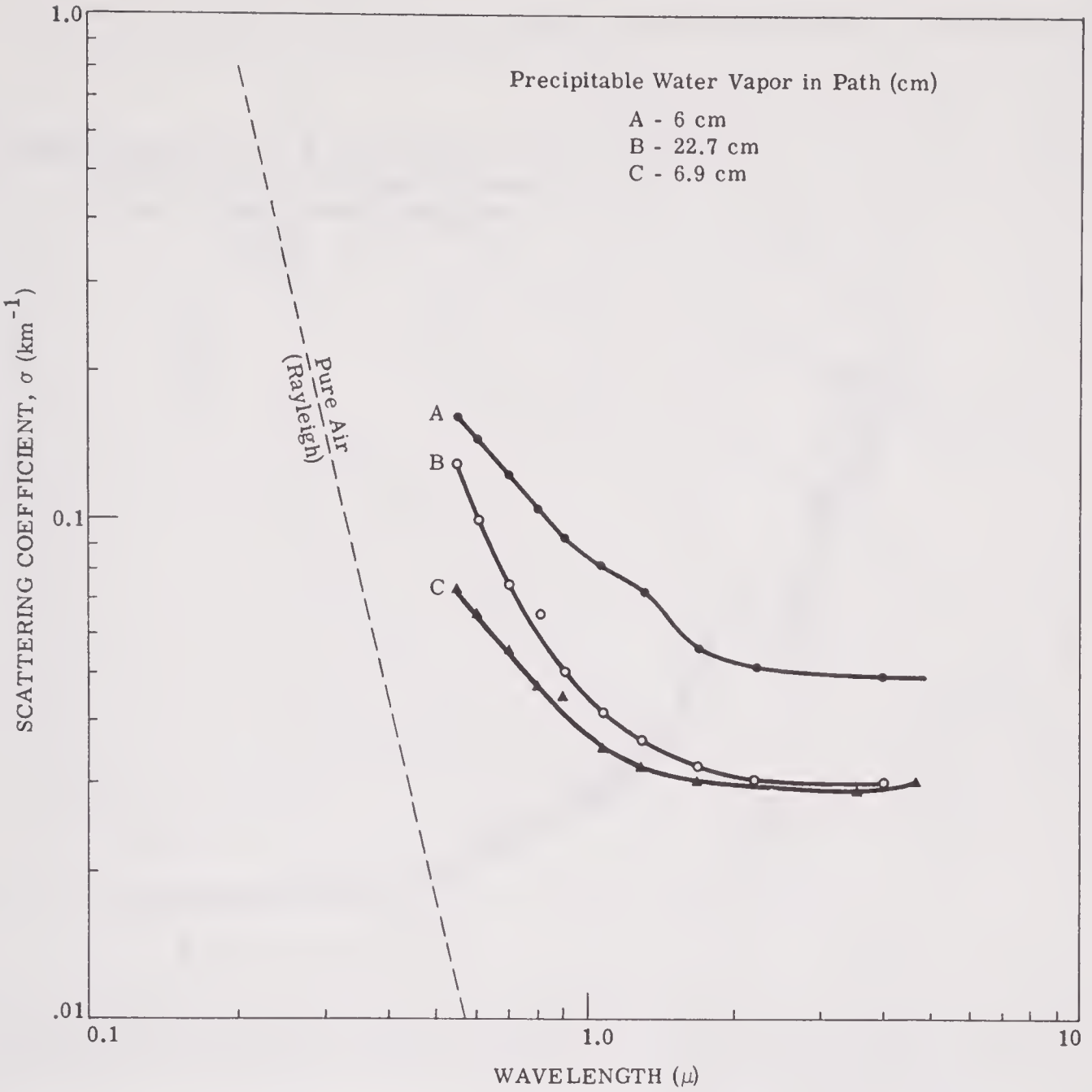


FIG. 6-30. Scattering coefficient measured over a 16.25-km sea-level, maritime path. Curve *B* represents a typical summer day condition [74].

In typical real atmospheres, the value of γ most recommended appears to be very nearly 1.3 [78]. It has been suggested [79] that the exponent is related to the meteorological range by the empirical form:

$$\gamma = 0.0585(MR)^{1/3} \tag{6-50}$$

This relationship, and various experimental data [80, 81] are shown in Fig. 6-37. The slopes of the data curves from Fig. 6-28 through 6-33 are also shown to give an indication of the slope in the infrared. The points are widely spread and significantly below the widely accepted value of 1.3. The smaller slopes indicated by the data are undoubtedly another manifestation of the contamination of the infrared windows by selective absorption.

6.6. Atmospheric Scintillation [82]

Scintillation (boil, shimmer, twinkle) is the variation in intensity or angular distribution of a collimated beam of radiation as a result of inhomogeneities in the atmospheric path through which the radiation passes. The following paragraphs briefly discuss these variations and their effects on optical systems. Detailed discussions of scintillation theory are given in [58, 59, 82-88].

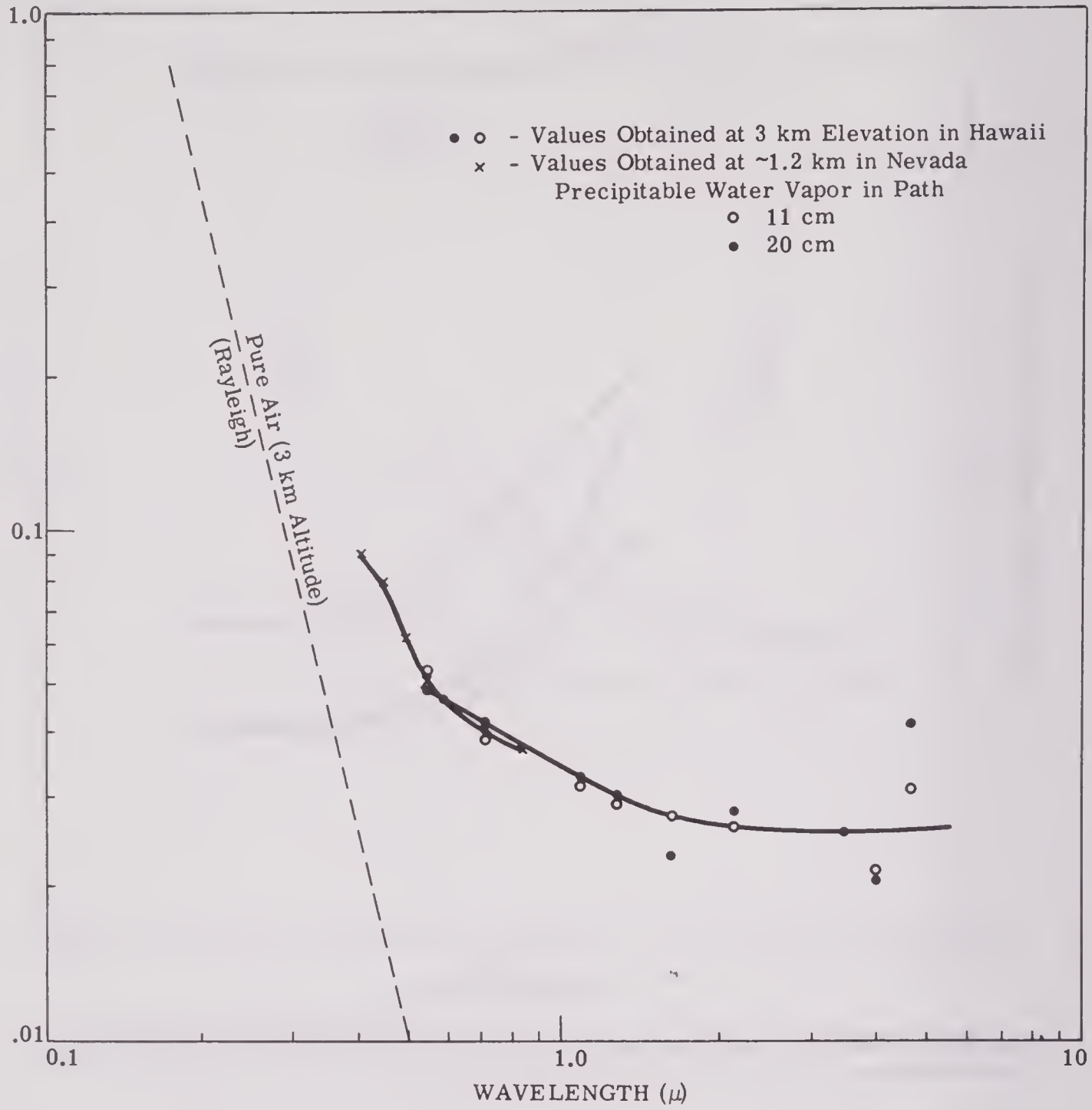


FIG. 6-31. Scattering coefficient measured over a 27.7-km horizontal path at 3-km elevation on the island of Hawaii [74] and the Nevada desert at ground level [75].

6.6.1. Inhomogeneities in the Atmosphere. Consider a three-dimensional pattern of refractive index inhomogeneities in the atmosphere in which there are discrete pattern elements that are somewhat spherical. Also, consider more or less continuous sheets of air differing in refractive index from the average, and having surface irregularities. Some of these surface irregularities will be somewhat hemispherical and can be treated as discrete elements also. In some cases, the interface on adjacent sides of a sheet will be plane, but not parallel.

Although the interface of these refractive elements may or may not be somewhat diffuse, the transit time for a collimated beam through a refractive element of lower than average refractive index will be shortened, while the transit time for the a beam passing through a refractive element of higher than average refractive index will be increased. Thus, there is a grouping of the "transit times" for the beams having passed through a small cubical volume containing a single refractive element. The nature or distribution of this grouping of transit times is comparable to the distribution of transit times for beams that have passed through a similar cubical volume containing an optical

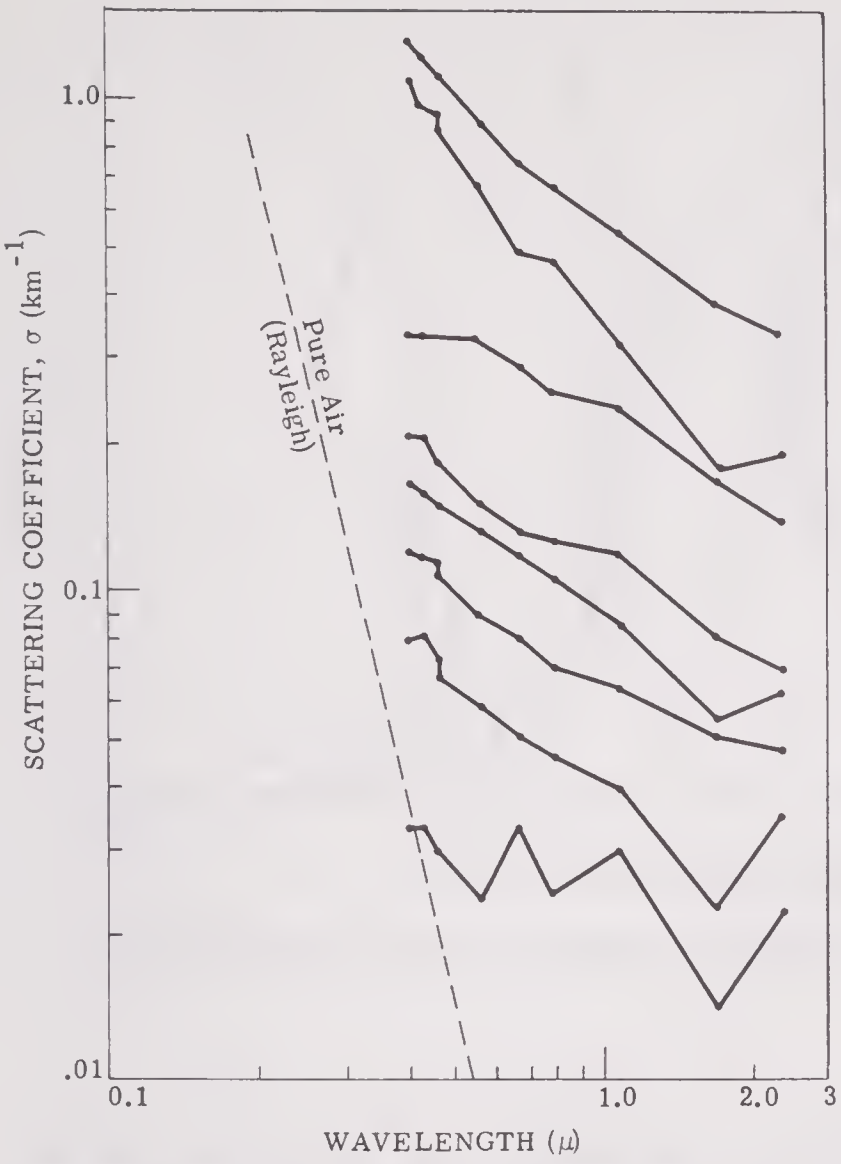


FIG. 6-32. Measured values of scattering coefficient as a function of wavelength at different times [76].

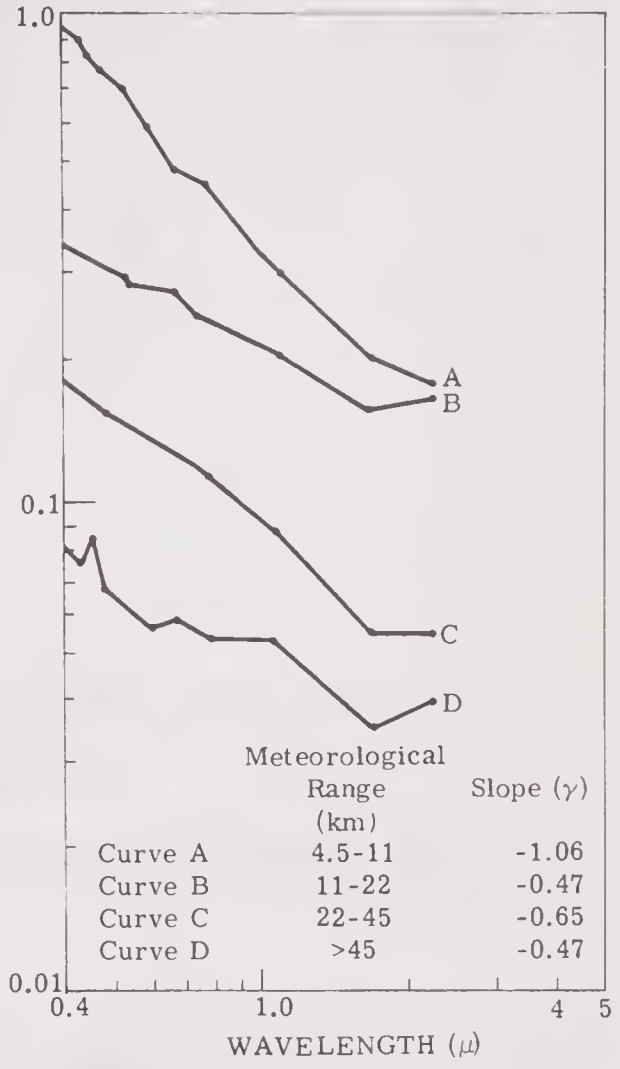


FIG. 6-33. Measured values of scattering coefficient grouped according to meterological range [76].

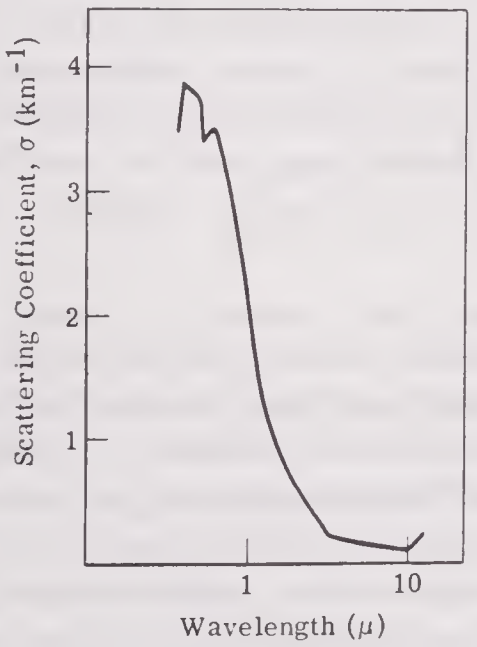


FIG. 6-34. Scattering coefficient for a haze [38].

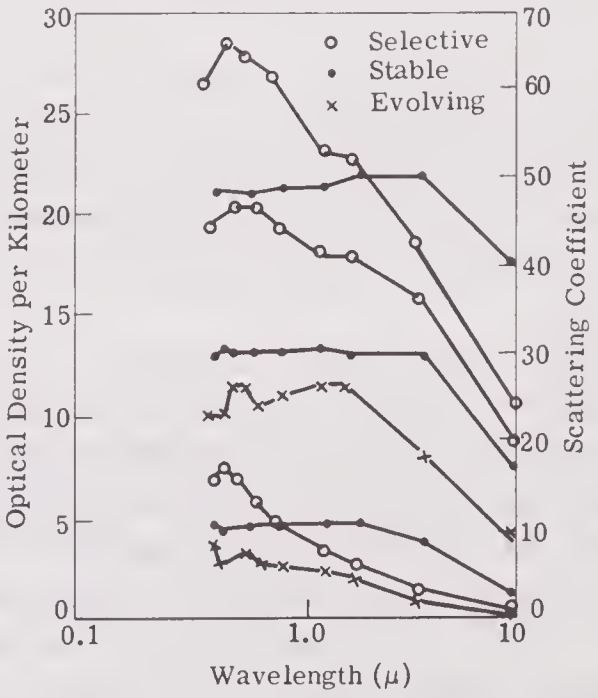


FIG. 6-35. Optical density of fogs [38].

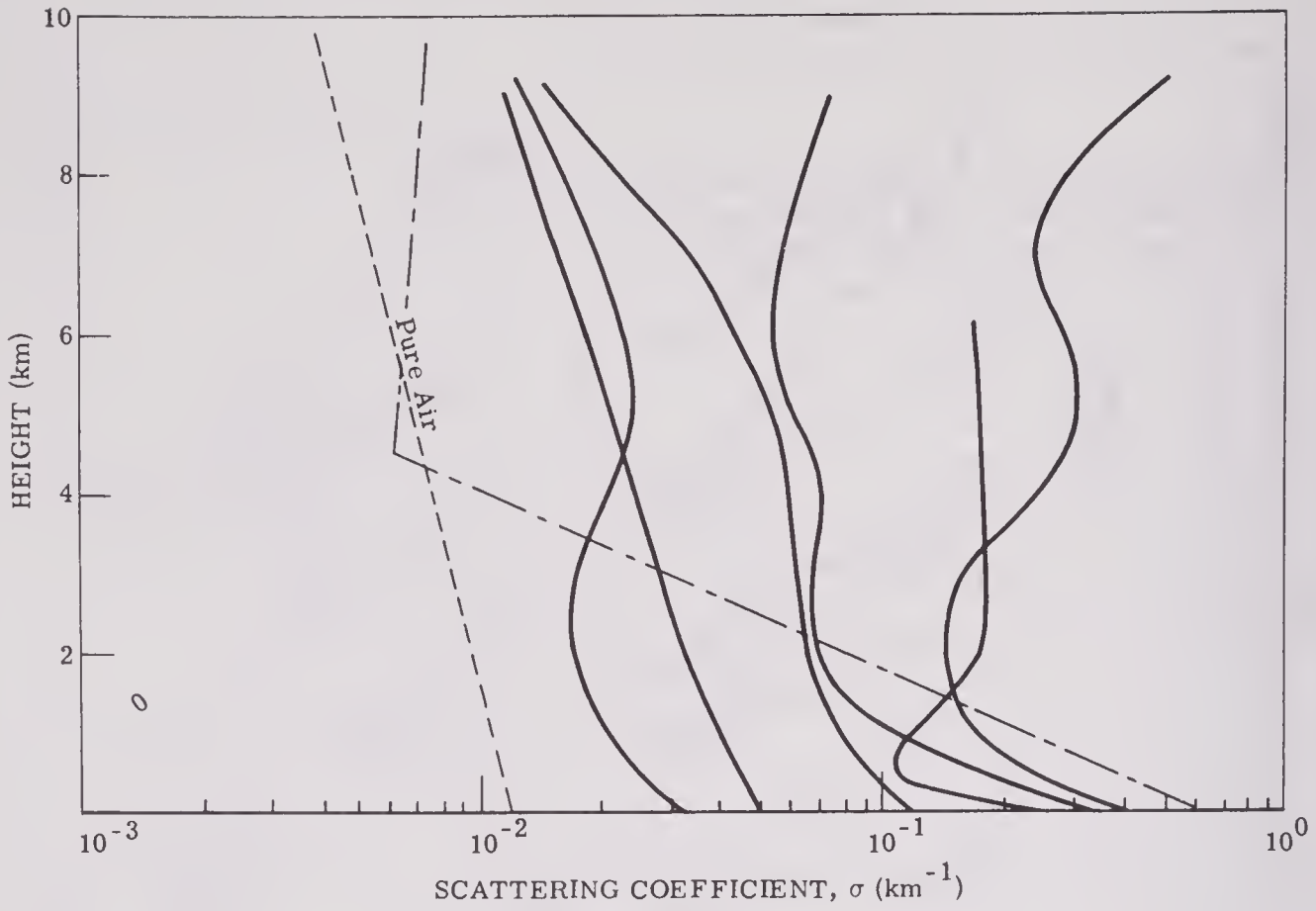


FIG. 6-36. Scattering coefficient as a function of height for a variety of conditions in clean air.

element such as a lens or prism. Therefore, the shapes of the wave fronts and the angular distributions of the beams are also comparable. To the extent that the refractive elements are in reality comparable to an optical element, they may be so treated analytically.

A single refractive element in the size range of 1 inch to 1 foot in diameter can cause an appreciable change in the angular distribution of radiation in a portion of the beam, resulting in a marked change in the radiation level at a more distant point in the beam. Further, the extent of the change in the angular distribution of the radiation in the beam is dependent upon the size, shape, and refractive index differential of the refractive element. The extent of change in the radiation level at a more distant point in the beam is also dependent, however, on the relative location of the refractive element in the optical path, as well as on the total path length. And in some cases where diffraction is predominant, the change in the received radiation level is also a function of wavelength.

6.6.2. Image Boil. Consider an optical system that is aimed at a small, distant source. A region of refractive index inhomogeneities causes nonuniformities of the angular orientation of rays within a beam of radiation, and these angular nonuniformities lead to intensity variations in the plane of the receiver aperture. It follows that these nonuniformities in the angular orientation of rays can also produce deviations in the distribution of radiation in the image plane of the optical system. As a result, the very small image produced under clear, homogeneous conditions can become spread out or shifted in position when the incident angular distribution of rays becomes spread out or shifted in direction. This effect has been observed photographically and visually, and is often referred to in the literature as "image boil."

If the rays incident on the optical system are all deviated a like amount and in the same direction, as though having passed through a prism, the image of the source

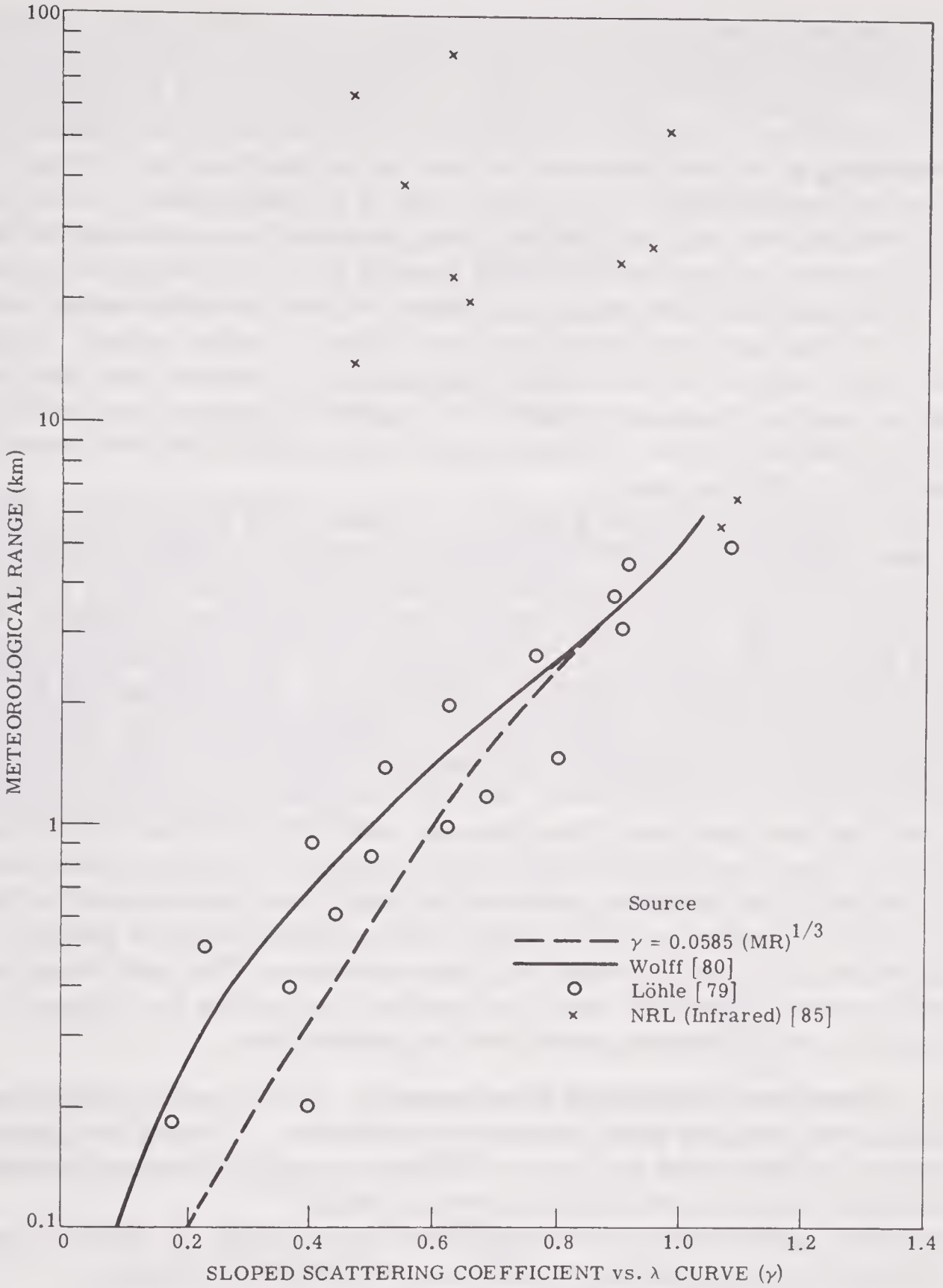


FIG. 6-37. Experimental values of slope of scattering coefficient curves.

retains its small size but is displaced from the optical axis. Displacements of stellar images of as much as 5 to 10 seconds of arc have been measured. Displacements of as much as 15 to 30 seconds of arc can be expected under severe ground-to-ground conditions. Because of the erratic motion associated with the observed displacement of the image, the term "image jitter" is often used for this effect.

6.6.3. Enlargement of the Image. To the extent that the rays incident on an optical system are either dispersed or converged symmetrically about the direction of the optical axis, the image of the small, distant source becomes enlarged. The same effect would result if a lens were placed on the optical axis. If the lens were positive, the beam would become more convergent and form a small image in front of the focal

plane. If the lens were negative, the beam would become more divergent and form a small image behind the focal plane. In both cases, the image in the focal plane would be spread. The amount of "image spread" is dependent on the broadening of the angular distribution of rays that are incident on the aperture of the optical system. This broadening is, in turn, dependent not only on the focal length of the lens, but on the size of the lens in relation to the aperture size of the optical system and on the distance between the lens and the aperture. Thus, if a large lens were placed in front of the optical system, only the center (or least deviated part) of the bundle of rays would enter. If the large lens were positive and backed off from the optical system, a larger portion of the angularly distributed rays would enter the optical system, increasing the size of the image in the focal plane. Conversely, if a negative lens were backed off from the aperture, a smaller portion of the angularly distributed rays would enter the optical system and the size of the focal plane image would decrease and become more the size of the undisturbed image.

The same changes in focal-plane image sizes would result if lenses of shorter and longer focal lengths were placed immediately in front of the optical system. Thus, there are two mechanisms for producing image-size changes that cannot be differentiated directly. There is a difference between the two mechanisms, however. When the focal length of the close-up lens is changed, there is no accompanying change in the total radiation in the optical system; when the large lens is backed off, there is a change in the total received radiation because a greater portion of the large positive lens will be effective in collecting radiated energy to concentrate onto the small receiver aperture as the lens is backed off more. Actually, the increase in total received radiated energy will continue until one of the following occurs: (1) all of the radiated energy falling on the lens is converged onto the receiver aperture, (2) the size of the converged beam is limited by the diffraction pattern of the lens, (3) the lens is unable to form an image. For the negative lens, more energy will be diverted from the aperture of the optical system as the lens is backed off a greater distance. The total energy in this case will continue to decrease until the negative lens reaches the midpoint of the optical path, or until diffraction around the lens predominates.

6.6.4. Atmospheric Scintillation Measurements. Data on stellar scintillation and scintillation over land and water are given in this section. The data are typical, and vast amounts of similar data are given in the references cited. Causes of scintillation and its effects on optical systems are discussed in Sec. 6.6.

Scintillation measurements can be expressed conveniently in terms of percent equivalent sine wave modulation per unit bandwidth, which is defined as

$$M(\%) = 1.414\overline{E}_f/\overline{E}_{dc} \tag{6-51}$$

where $\overline{E}_f$ = average noise voltage at frequency f
$\overline{E}_{dc}$ = average dc level of the signal

Often, the average percent equivalent sine wave modulation, $\overline{M}(\%)$, is plotted to keep the figures for all frequency ranges at about the same value. The average value in a given frequency range can be converted to root mean square deviation of the signal in that range by multiplying it by the square root of the frequency interval covered.

The frequency range of scintillation components is approximately 2.5 to 450 cps. To simplify scintillation measurements, this frequency range can be broken down into subranges, *e.g.*, 2.5 to 10.5 cps, 10 to 50 cps, and 50 to 450 cps, and expressed as $\overline{M}(\%)_{2.5-10.5}$, $\overline{M}(\%)_{10-50}$, and $\overline{M}(\%)_{50-450}$.

Scintillation modulus (mod), which is sometimes used to define the shape of the curve, is the ratio of the average percent equivalent sine wave modulation in the mid-frequency or high-frequency range to that of the low-frequency range.

$$\text{mod}_{10-50} = \frac{\overline{M}(\%)_{10-50}}{\overline{M}(\%)_{2.5-10.5}}$$

$$\text{mod}_{50-450} = \frac{\overline{M}(\%)_{50-450}}{\overline{M}(\%)_{2.5-10.5}} \qquad (6\text{-}52)$$

6.6.4.1. *Stellar Scintillation* [83, 84]. Figure 6-38 shows a typical plot of percent equivalent sine wave modulation, $M(\%)$, *versus* frequency (z = zenith angle). The shape of the curve can be characterized by the rollover and crossover frequencies. The rollover frequency is the frequency at which an average line drawn through the low-frequency components intersects with an average line drawn through the mid- and high-frequency components. The crossover frequency is the frequency at which the high-frequency average line intersects the zero axis. Determination of the rollover and crossover frequencies depends, to a certain extent, on the manner in which the scintillation curves are plotted. Figure 6-38, and subsequent stellar scintillation curves, are based on semilog plots.

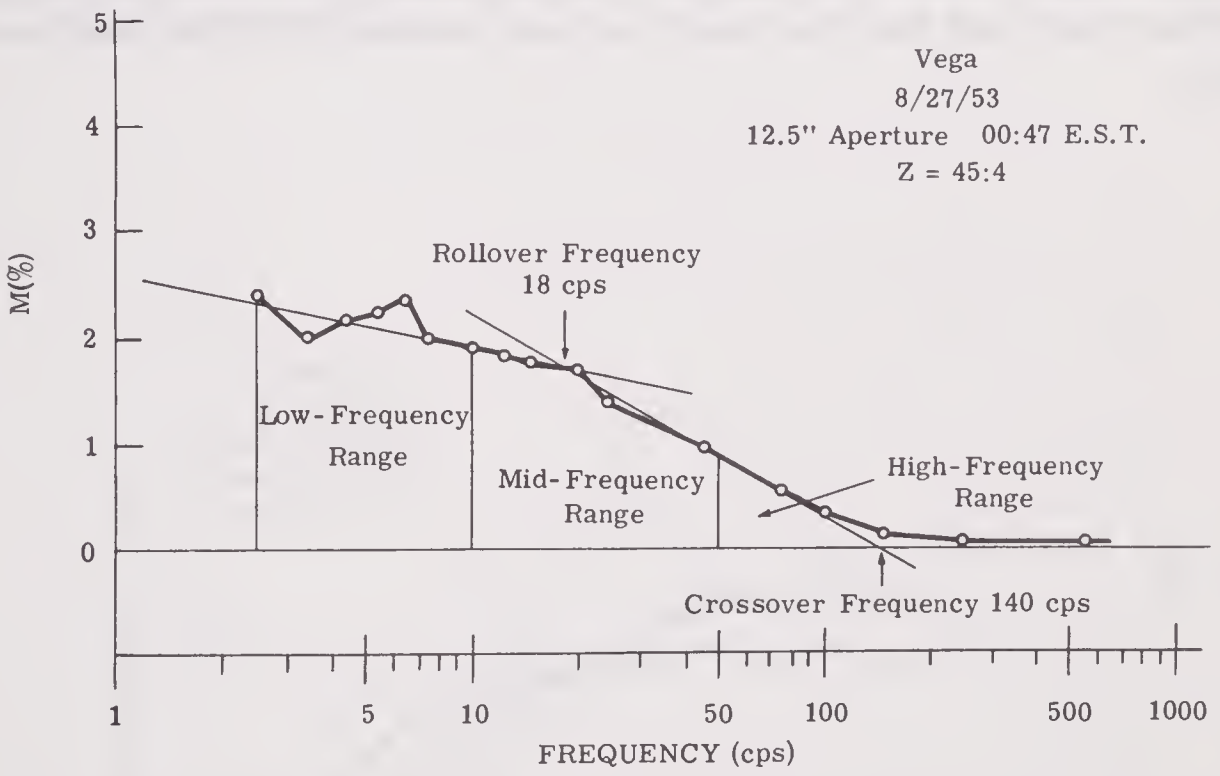


FIG. 6-38. A representative frequency analysis of stellar scintillation [59].

DIURNAL AND SEASONAL VARIATIONS IN STELLAR SCINTILLATION. Figure 6-39 is a comparison of daytime and nighttime scintillation. The shape of the daytime and nighttime curves are very similar, and the amount of scintillation is only slightly greater during the daytime.

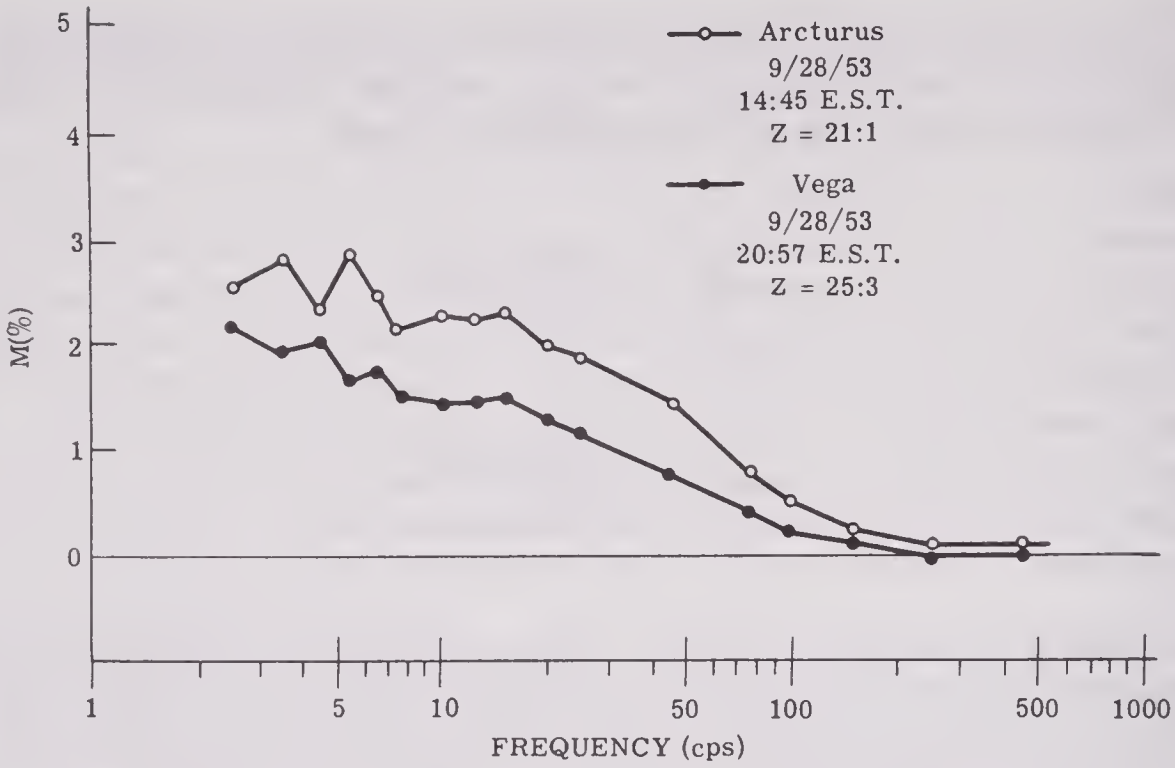


FIG. 6-39. Comparison of daytime to nighttime stellar scintillation [59].

Figure 6-40 shows the typical seasonal variation of stellar scintillation. Total scintillation is generally greater in the winter, but the low-frequency components are greater during the summer.

EFFECTS OF UPPER-AIR WINDS ON STELLAR SCINTILLATION. Stellar scintillation appears to be virtually independent of surface weather conditions. At the 200-mb level (approximately 40,000 ft), however, a correlation exists between the shape of the scintillation curve and the wind speed, and also between slit orientation and wind direction.

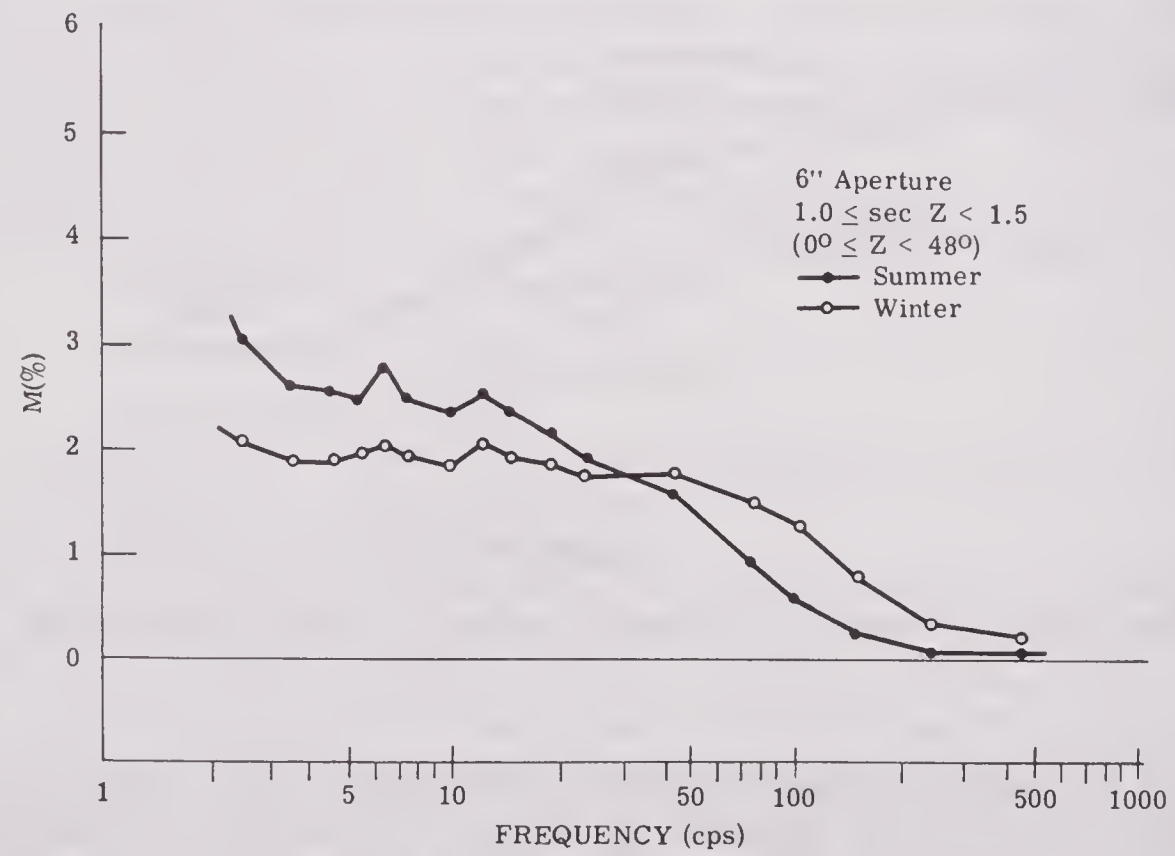


FIG. 6-40. Comparison of winter to summer stellar scintillation [59].

Figure 6-41 shows the relationship between wind speed and the shape of the scintillation curve at the 200-mb level. Wind speeds are estimated to be accurate within ±10 knots. The data in Fig. 6-41 were obtained using a 12.5-in. aperture, but the correlation is valid for smaller apertures also.

Figure 6-42 illustrates the correlation between slit orientation and wind direction. The dashed lines represent deviations of ±20°, which are estimates of the probable error in reading wind directions from plotted weather maps.

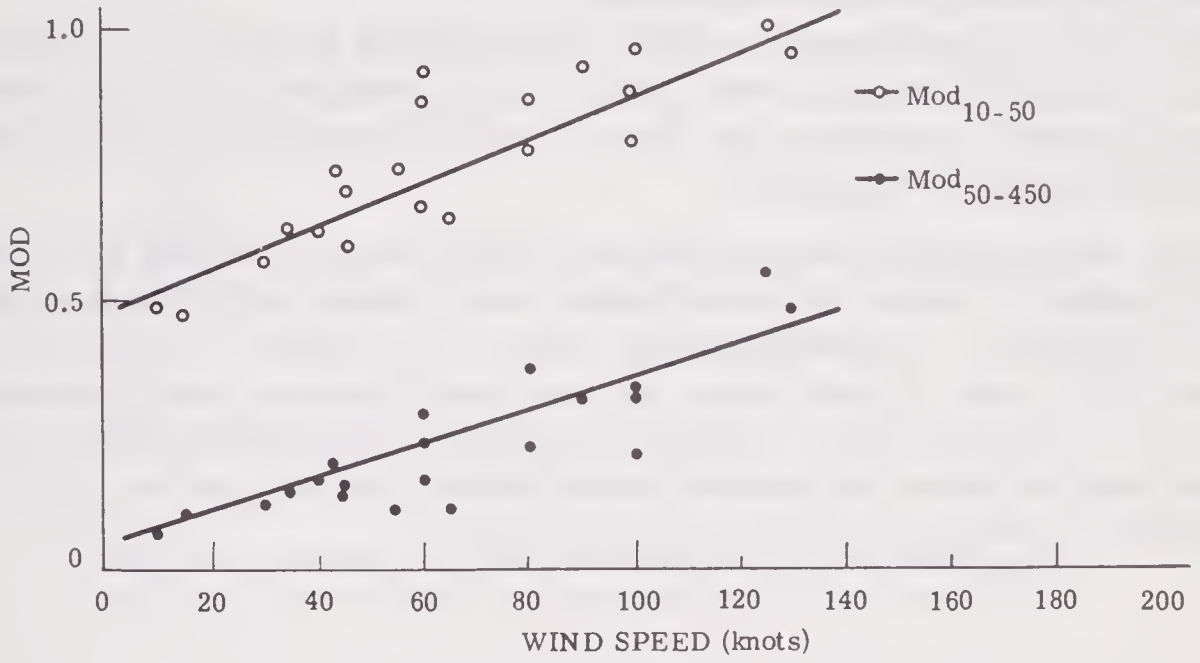


FIG. 6-41. Relation of scintillation moduli to wind speed at 200-mb level for a 12.5-in. aperture [59].

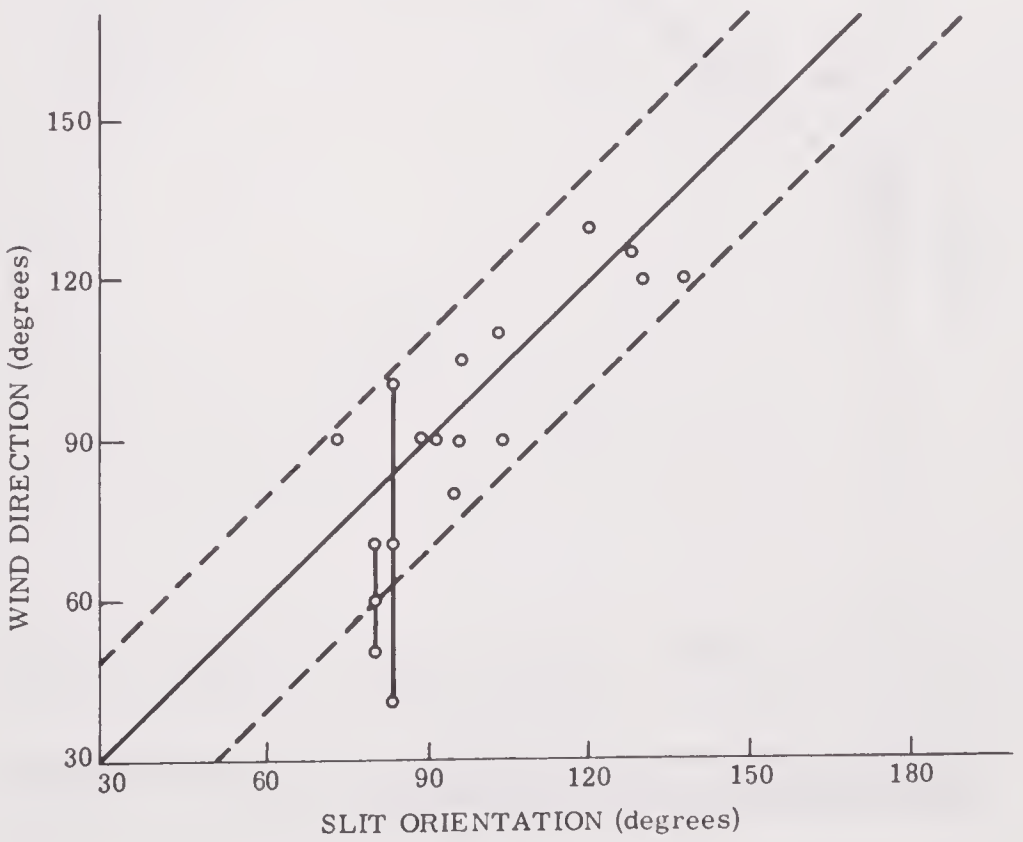


FIG. 6-42. Correlation of slit position for minimum 200-cps scintillation component with wind direction at 200-mb level [59].

Effects of Zenith Distance on Stellar Scintillation. Stellar scintillation is greater near the horizon than it is overhead. It is difficult to precisely describe this effect, however, because the manner in which scintillation varies with the secant of the zenith angle depends greatly upon aperture size, the frequency range under consideration, and meteorological factors. The latter complicate the determination of zenith distance considerably.

Figures 6-43 and 6-44 show $\overline{M}(\%)_{2.5-10.5}$ and $\overline{M}(\%)_{total}$ *versus* log sec Z, respectively, for a 12-in. aperture. The same data for a 3-in. aperture are shown in Fig. 6-45 and 6-46. No meaningful data are available for the mid-frequency scintillation range (10 to 50 cps) and the high-frequency range (50 to 450 cps) because of the very strong meteorological affects at these frequencies.

Table 6-5 gives approximate functional relationships between $\overline{M}(\%)$ and sec Z of the form $a(\sec Z)^n$ for various aperture sizes. The values given in the table represent a first-order attempt to describe the effect of zenith distance on scintillation in that wind-velocity effects are neglected.

Effects of Aperture Size on Stellar Scintillation. Figure 6-47 shows the general manner in which the scintillation curve changes with aperture size. The ratio of scintillation at various aperture sizes to scintillation for a 12.5-in. aperture is given in Fig. 6-48. In both curves, the relationship between scintillation and aperture is for low wind velocity. The combined effects of both wind velocity and zenith distance upon the ratio of the amount of scintillation for a 3-in. aperture and the amount for a 12.5-in. aperture is shown in Fig. 6-49 and 6-50.

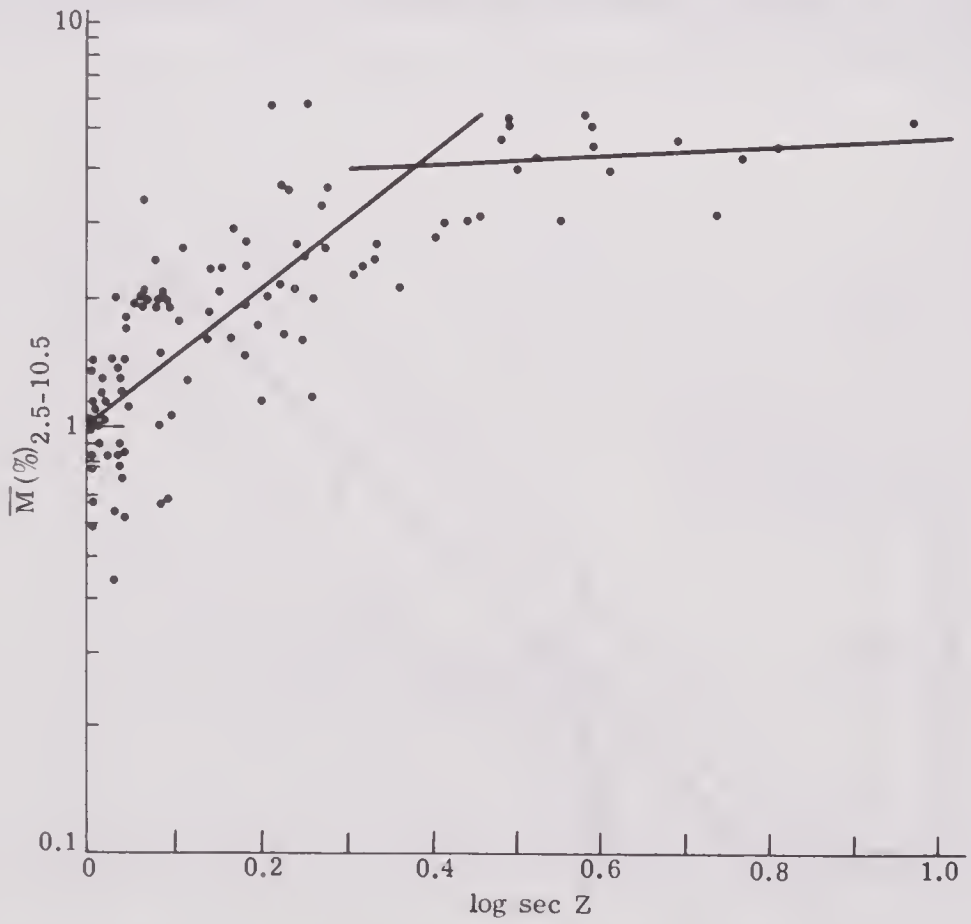


Fig. 6-43. Variation of low-frequency stellar scintillation with zenith distance for 12.5-in. aperture [59].

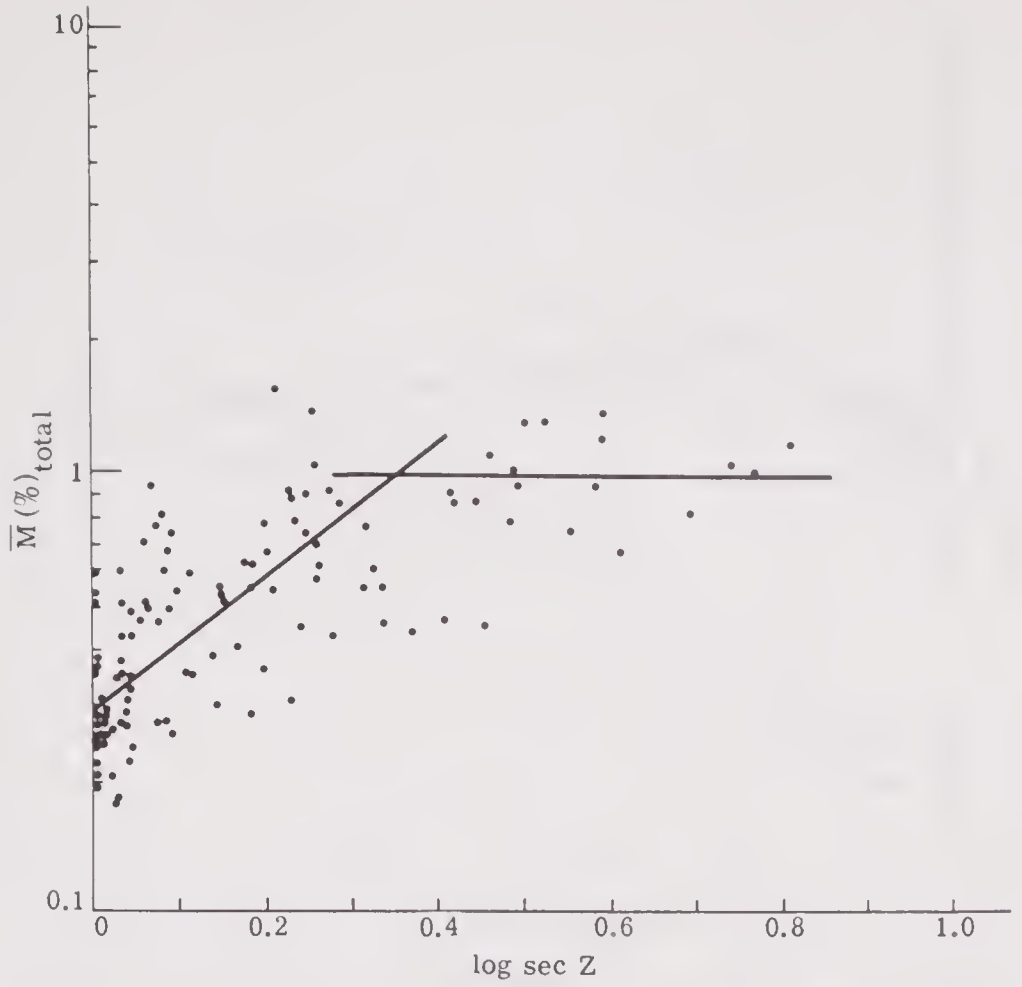


FIG. 6-44. Variation of stellar scintillation over all frequencies with zenith distance for 12.5-in. aperture [59].

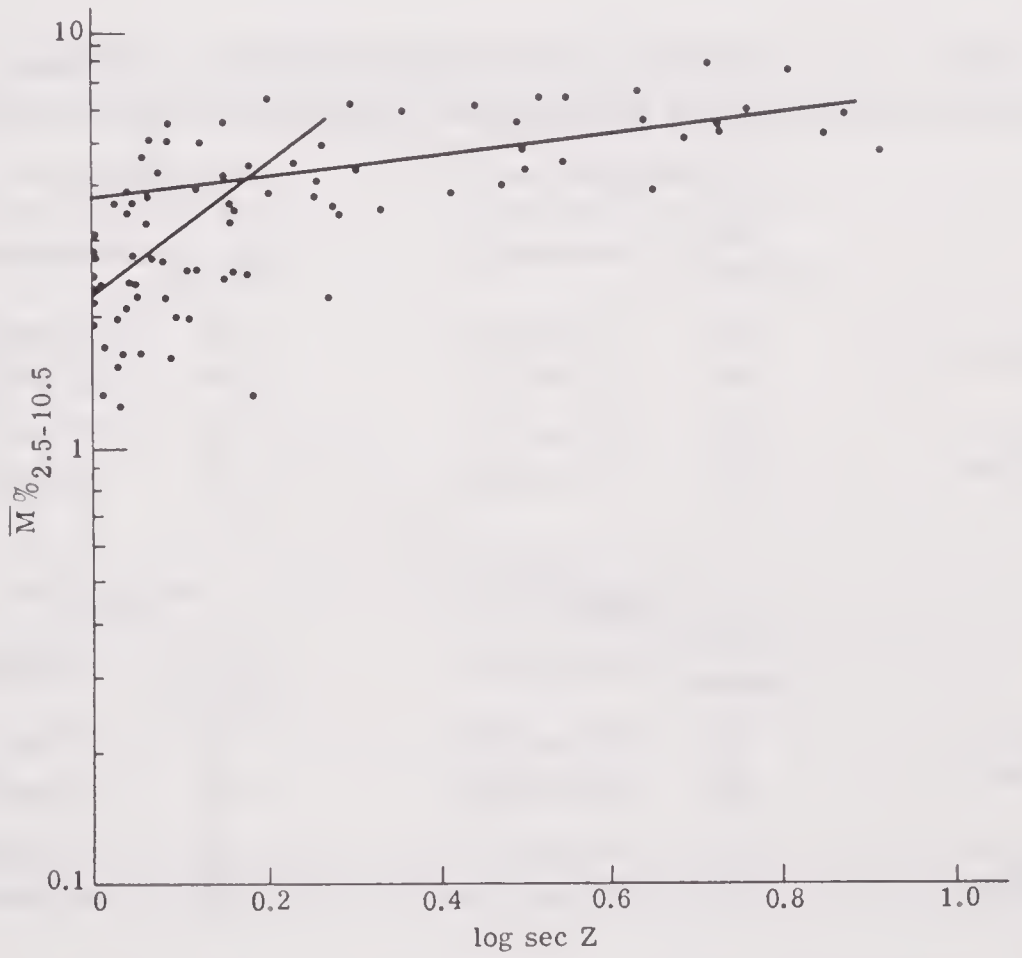


FIG. 6-45. Variation of low-frequency stellar scintillation with zenith distance for 3-in. aperture [59].

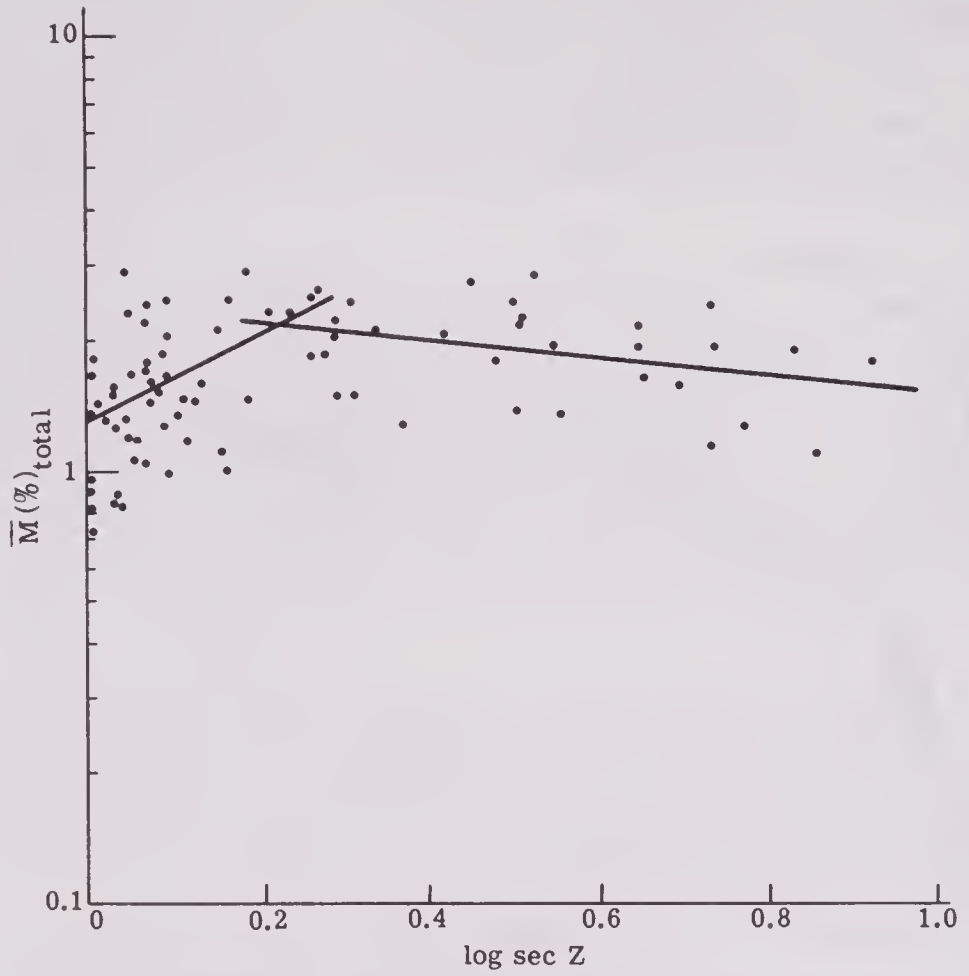


FIG. 6-46. Variation of stellar scintillation over all frequencies with zenith distance for 3-in. aperture [59].

TABLE 6-5. FUNCTIONAL RELATIONSHIPS BETWEEN AVERAGE EQUIVALENT SINE WAVE MODULATION AND ZENITH DISTANCE [83].

	12-*in.* *Aperture*		6-*in.* *Aperture*	
	Z (*Degrees*)	*Relationship*	*Z* (*Degrees*)	*Relationship*
$\overline{M}(\%)_{2.5-10.5}$	65	1.0 (sec Z)$^{1.6}$	60	1.5 (sec Z)$^{1.6}$
	65	3.6 (sec Z)$^{0.1}$	60	3.6 (sec Z)$^{0.3}$
$\overline{M}(\%)_{total}$	63	0.3 (sec Z)$^{1.5}$	58	0.7 (sec Z)$^{1.2}$
	63	1.0 (sec Z)$^{0.0}$	58	1.4 (sec Z)$^{0.0}$

	3-*in.* *Aperture*		1-*in.* *Aperture*	
	Z (*Degrees*)	*Relationship*	*Z* (*Degrees*)	*Relationship*
$\overline{M}(\%)_{2.5\text{-}10.5}$	49	2.3 (sec Z)$^{1.5}$	48	2.7 (sec Z)$^{1.6}$
	49	3.7 (sec Z)$^{0.3}$	48	4.5 (sec Z)$^{0.3}$
$\overline{M}(\%)_{total}$	53	1.3 (sec Z)$^{1.0}$	52	1.7 (sec Z)$^{0.9}$
	53	2.4 (sec Z)$^{-0.2}$	52	3.0 (sec Z)$^{-0.2}$

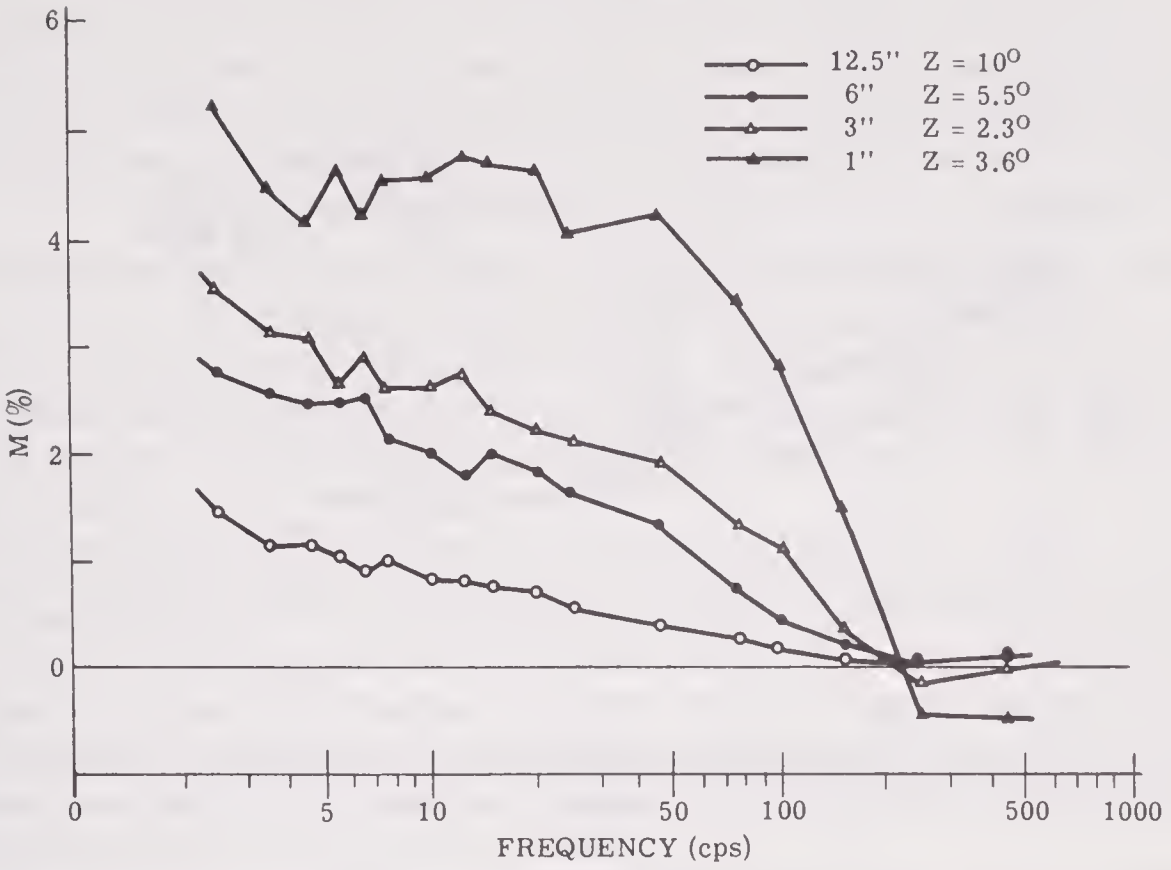


FIG. 6-47. Variation of scintillation curve with aperture [59].

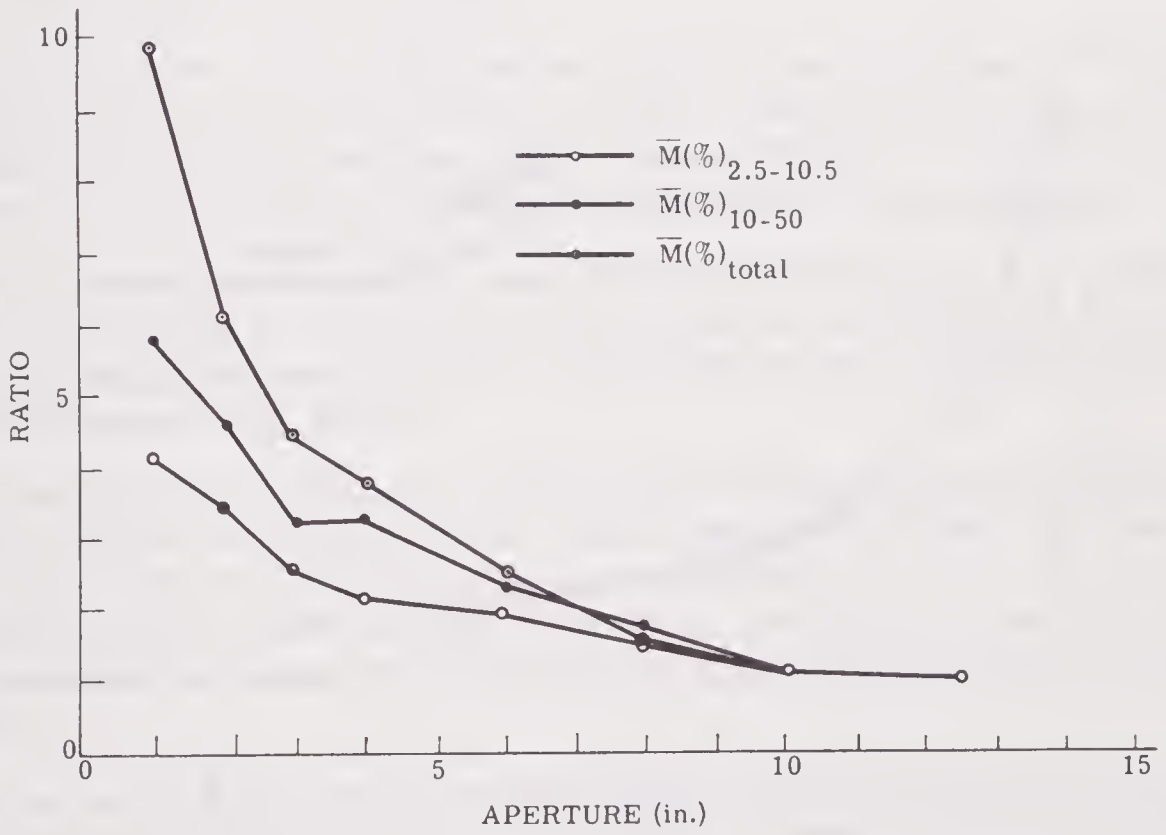


FIG. 6-48. Relationship between amount of stellar scintillation for various aperture sizes expressed in terms of scintillation from 12.5-in. aperture [59].

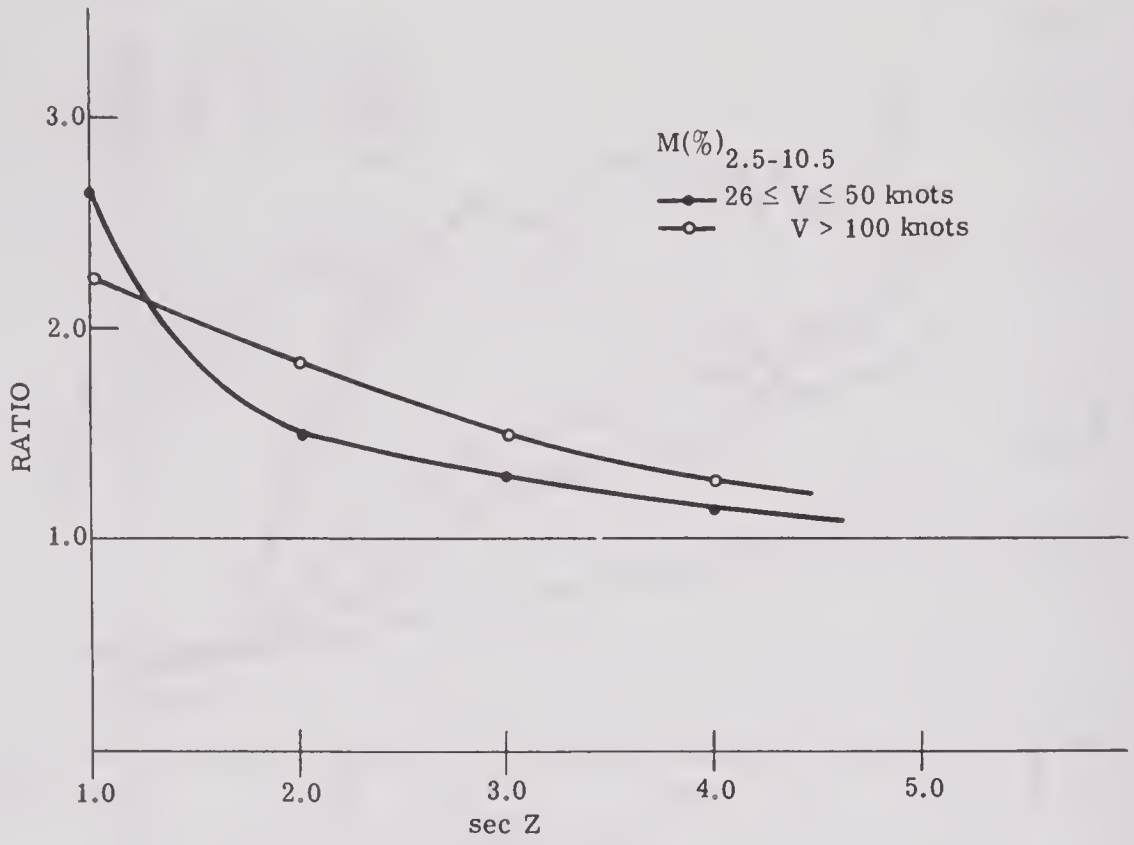


FIG. 6-49. Effect of zenith distance upon ratio of the low-frequency scintillation components for 3-in. aperture to a 12-in. aperture for different wind velocities at 200 mb [59].

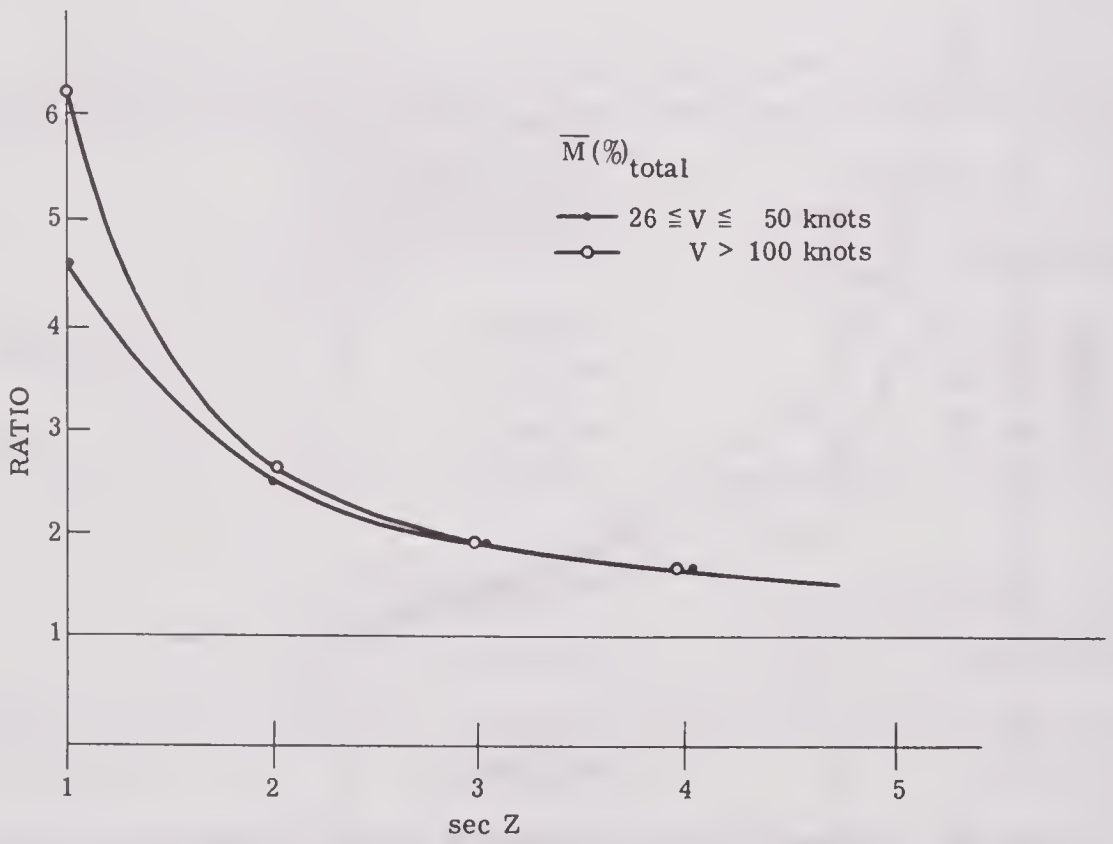


FIG. 6-50. Effect of zenith distance upon the ratio of the scintillation taken over all frequencies for a 3-in. to a 12.5-in. aperture for different wind velocities at 200 mb [59].

6.6.4.2. *Scintillation Over Water* [85, 86].

Seasonal Variations. Figure 6-51 shows seasonal variations of atmospheric scintillation over water along a 17,750-yd path. The data were obtained with a 12-in. searchlight mounted 30 ft above the water and a 24-in. optical receiver mounted 109 ft above the water. The values of percent equivalent sine wave modulation, $M(\%)$, shown in Fig. 6-51, are averages for each month. The air temperatures are also average values for each month, based on readings obtained at both ends of the transmission path.

Effects of Receiver Collector Area. Figure 6-52 shows representative frequency spectrum curves for four different receiver collector areas. The manner in which scintillation varies with receiver collector area is shown in Fig. 6-53. The data are based on measurements over water along a 4400-yd path. The transmitting source was mounted approximately 18 ft above the water and the receiver approximately 109 ft above the water.

Effects of Source Area. Figure 6-54 shows representative frequency-spectrum curves for four different source areas. The effective area of each lamp source is approximately 2.8 $in.^2$, and the angle subtended by each lamp is approximately 2.2 seconds of arc. Figure 6-55 shows the manner in which total scintillation varies with source area. The data in both illustrations were measured over water along a 4400-yd path. The transmitting source was mounted approximately 18 ft above the water and the receiver approximately 109 ft above the water.

The amount of scintillation decreases as the number of source lamps (and, consequently, the area of the source) increase. However, the physical distribution and orientation of the source lamps with respect to each other appear to have little effect on the amount of scintillation.

6.6.4.3. *Scintillation Over Land* [87].

Diurnal Variations. Figure 6-56 shows typical diurnal variations in the amount of scintillation over a grass surface on a clear day. The data were determined at the mid-latitudes near the solar equinox. This pattern of daily variations is subject to many influences, but it will retain the general relationship to solar elevation.

Effects of Temperature and Wind on Scintillation Over Land. In general, the amount of scintillation is proportional to the absolute magnitude of temperature change with height, although its characteristics depend upon whether temperature increases or decreases with height.

(*a*) For temperature decreasing markedly with height (an unstable atmosphere), the typical condition on a cloudless day (except over snow surfaces), scintillation increases rapidly with increases in temperature gradient but slowly with increase in wind speed for winds less than 10 mph. Therefore, under unstable conditions, scintillation is dependent on temperature gradient to a greater degree than on wind speed (Fig. 6-57*a*).

(*b*) For temperature increasing markedly with height (a stable atmosphere), the typical condition on a cloudless night, scintillation increases slowly with increases in temperature gradient but rapidly with increases in wind speed for winds up to 3 or 4 mph. Therefore, under stable conditions scintillation is dependent to a greater degree on wind speed than on the temperature gradient (Fig. 6-57*b*).

(*c*) For adiabatic or near-adiabatic conditions, a decrease in temperature of about 0.01°C per meter (a neutral atmosphere), scintillation is at a minimum or is absent, irrespective of wind.

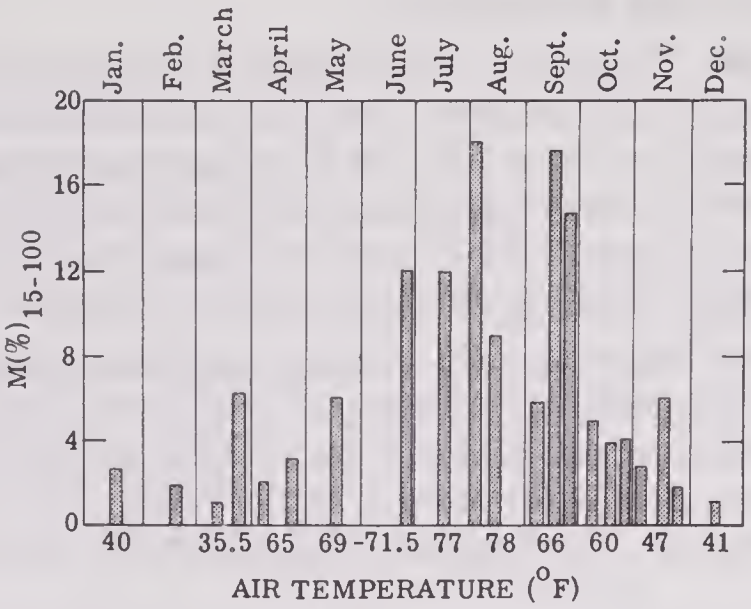


FIG. 6-51. Seasonal variation of atmospheric scintillation over a water path [62].

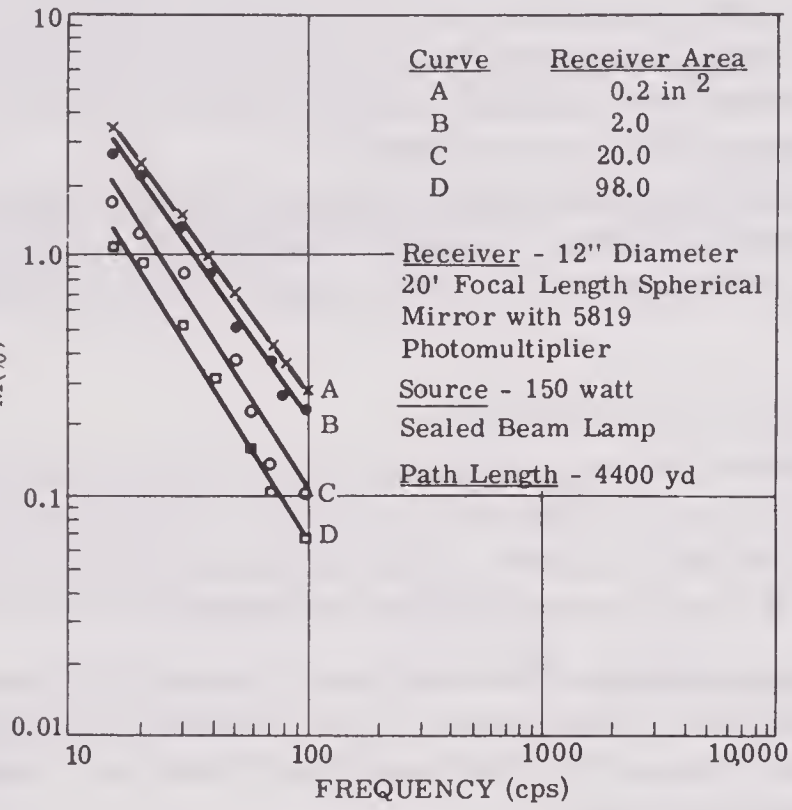


FIG. 6-52. Scintillation *vs.* frequency measured over water [62].

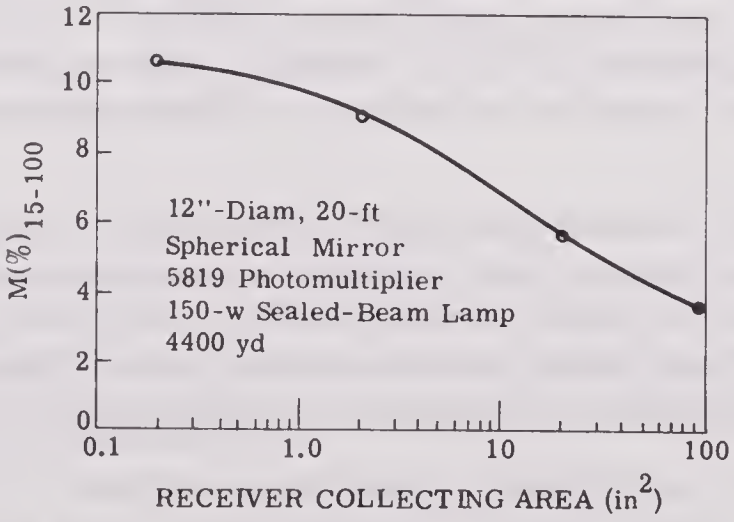


FIG. 6-53. Scintillation *vs.* receiver collector area measured over water [62].

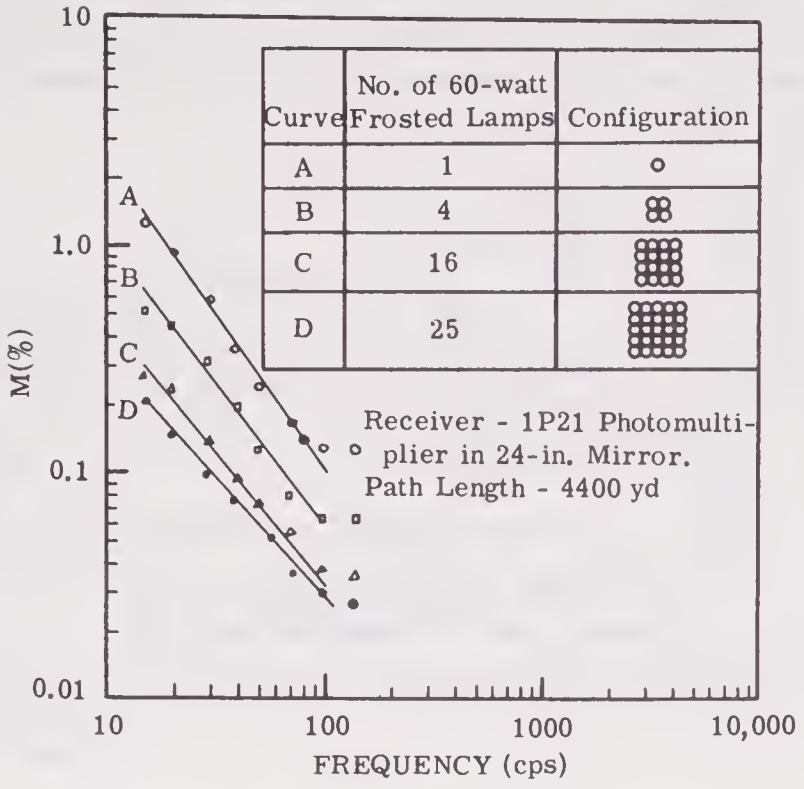


FIG. 6-54. Amount of scintillation over water *vs.* frequency for various source areas [62].

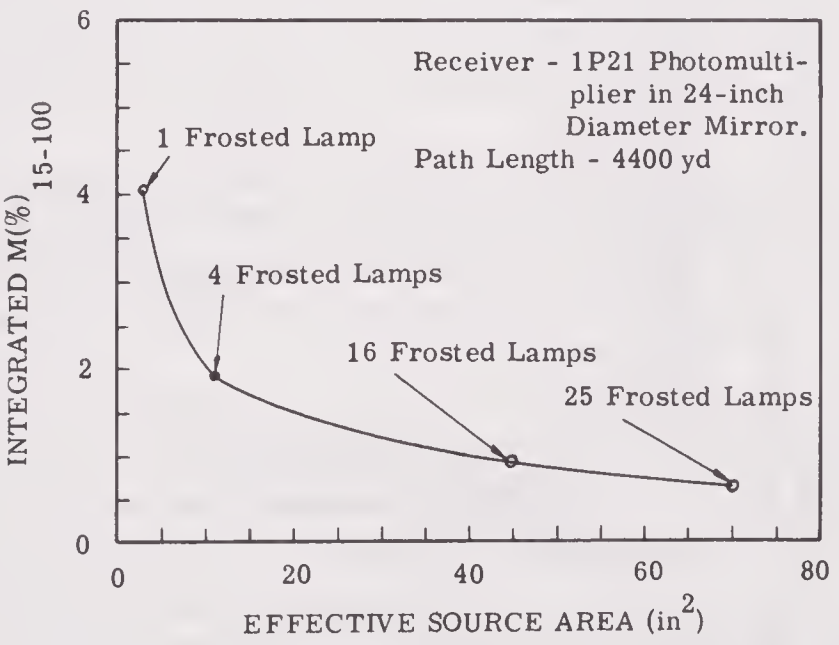


FIG. 6-55. Scintillation *vs.* source area measured over water [62].

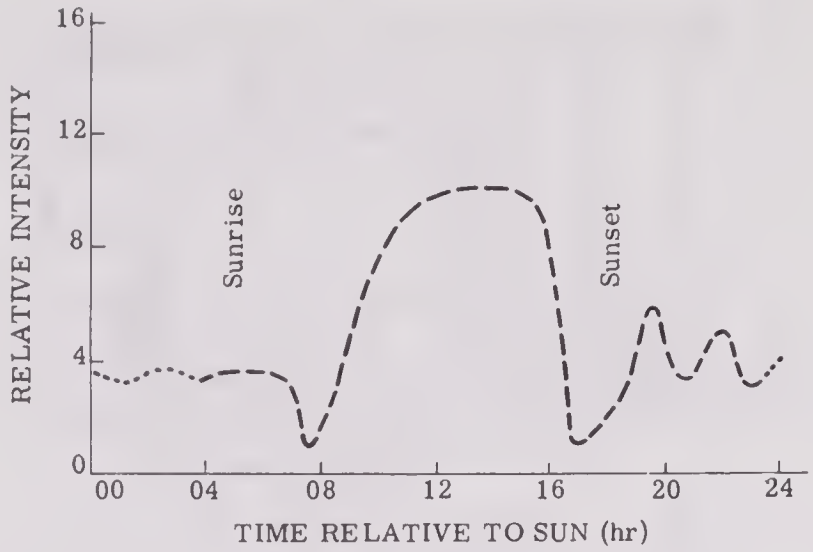


FIG. 6-56. Diurnal cycle of scintillation over a grass surface [83].

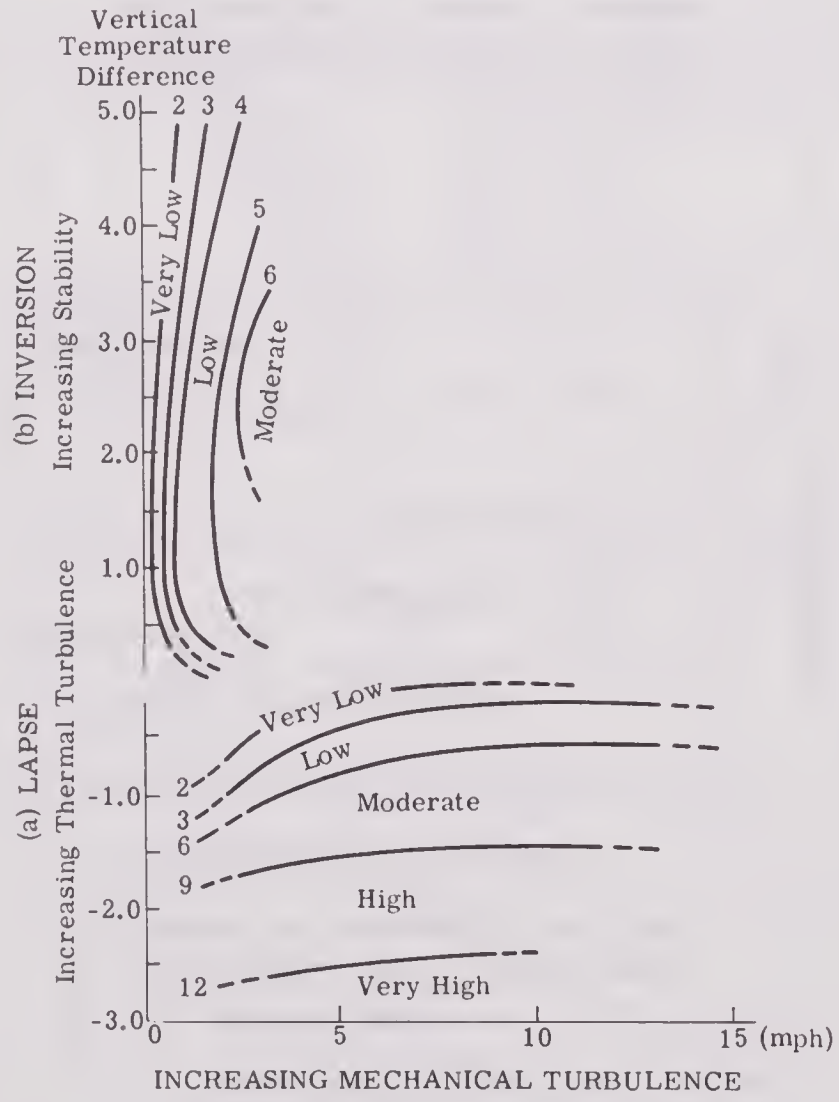


FIG. 6-57. Percent modulation as a function of temperature difference and wind speed (temperature difference between 0.5-m height; wind speed measured at 2-m height) [83].

Figure 6-58 shows the effects of an unstable, a stable, and a neutral atmosphere on the frequency of atmospheric scintillation. Three distinct conditions are illustrated: (1) midafternoon, with strong lapse conditions following the time of maximum heating of the ground; (2) the sunset period, during which the air is close to neutral equilibrium; (3) at night, after an inversion has formed because of surface cooling.

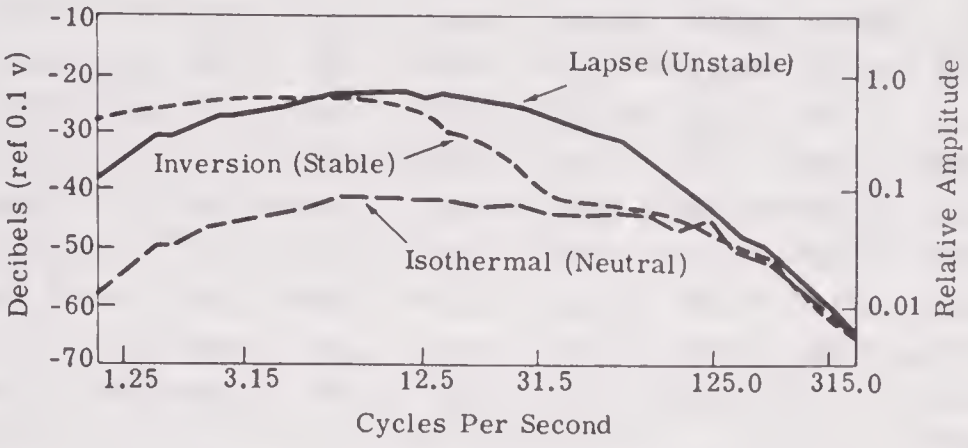


FIG. 6-58. Effects of a stable, unstable, and neutral atmosphere on atmospheric scintillation over a grass surface [83].

6.7. Solar Spectrum Measurements

A solar spectrum is a transmission spectrum of the earth's atmosphere in which the sun is used as the source of radiation. Many Fraunhofer lines are observed in the visible and ultraviolet portions of the solar spectrum. In the infrared, however, very few Fraunhofer lines are present and the sun emits approximately as a uniform 6000°K blackbody. Spectral details observed in the infrared solar spectrum are almost entirely due to the absorption of solar radiation by the molecules present in the earth's atmosphere.

The length of an absorption path through the atmosphere is dependent upon the elevation of the sun in the sky. Thus, when the sun is at the zenith, the solar radiation traverses one air mass of atmosphere. At any other angle from the zenith, called the solar altitude or zenith distance, the radiation traverses longer paths through the atmosphere.

Table 6-6 gives the equivalent air mass for a sea-level observer as the solar altitude varies from 0° to 90° [89]. To obtain the same type of table for an observer at a different reference altitude, multiply the values in Table 6-6 by the ratio of the pressure at the new reference level to that at sea level. This approximation holds well up to 85° angles and 100-km altitude.

Figure 6-59 shows a low-resolution solar spectrum for the region from 1 to approximately 15 μ. The other curves show the position and approximate relative intensities of the infrared absorption bands for various molecules in the atmosphere.

A number of high-resolution measurements of the solar spectrum have been made [90-103]. Figures 6-60 through 6-74 show the spectrum from approximately 2.80 to 14.2 μ [90-93]. The measurements were made from the Jungfraujoch, Switzerland, at an altitude of approximately 12,000 ft. Figures in the references contain more detail.

In the region from 2.80 to 3.15 μ (Fig. 6-60) the absorption is due mainly to H_2O, although lines of a band 13_{CO_2}, centered at 2.834 μ, $\nu_1 + \nu_3$ of N_2O, centered at 2.87 μ, and $2\nu_2 + \nu_3$ of N_2O, centered at 2.97 μ, are present. In the region of 3.15 to 3.50 μ (Fig. 6-61), the strong fundamental band ν_3 of CH_4, centered at 3.31 μ, is present. The remaining absorption is due to the weaker $2\nu_2$ band of H_2O.

TABLE 6-6. EQUIVALENT AIR MASSES FOR SOLAR ALTITUDES 0° TO 90°[a]

	0°	1°	2°	3°	4°	5°	6°	7°	8°	9°
0°	—	26.96	19.79	15.36	12.44	10.40	8.90	7.77	6.88	6.18
10°	5.60	5.12	4.72	4.37	4.08	3.82	3.59	3.39	3.21	3.05
20°	2.90	2.77	2.65	2.55	2.45	2.36	2.27	2.20	2.12	2.06
30°	2.00	1.94	1.88	1.83	1.78	1.74	1.70	1.66	1.62	1.59
40°	1.55	1.52	1.49	1.46	1.44	1.41	1.39	1.37	1.34	1.32
50°	1.30	1.28	1.27	1.25	1.24	1.22	1.20	1.19	1.18	1.17
60°	1.15	1.14	1.13	1.12	1.11	1.10	1.09	1.09	1.08	1.07
70°	1.06	1.06	1.05	1.05	1.04	1.04	1.03	1.03	1.02	1.02
80°	1.02	1.01	1.01	1.01	1.01	1.00	1.00	1.00	1.00	1.00
90°	1.00	—	—	—	—	—	—	—	—	—

[a]Entries in the table are the air masses for the angles indicated down the left and across the top. For example the air mass for 22° elevation is 2.65.

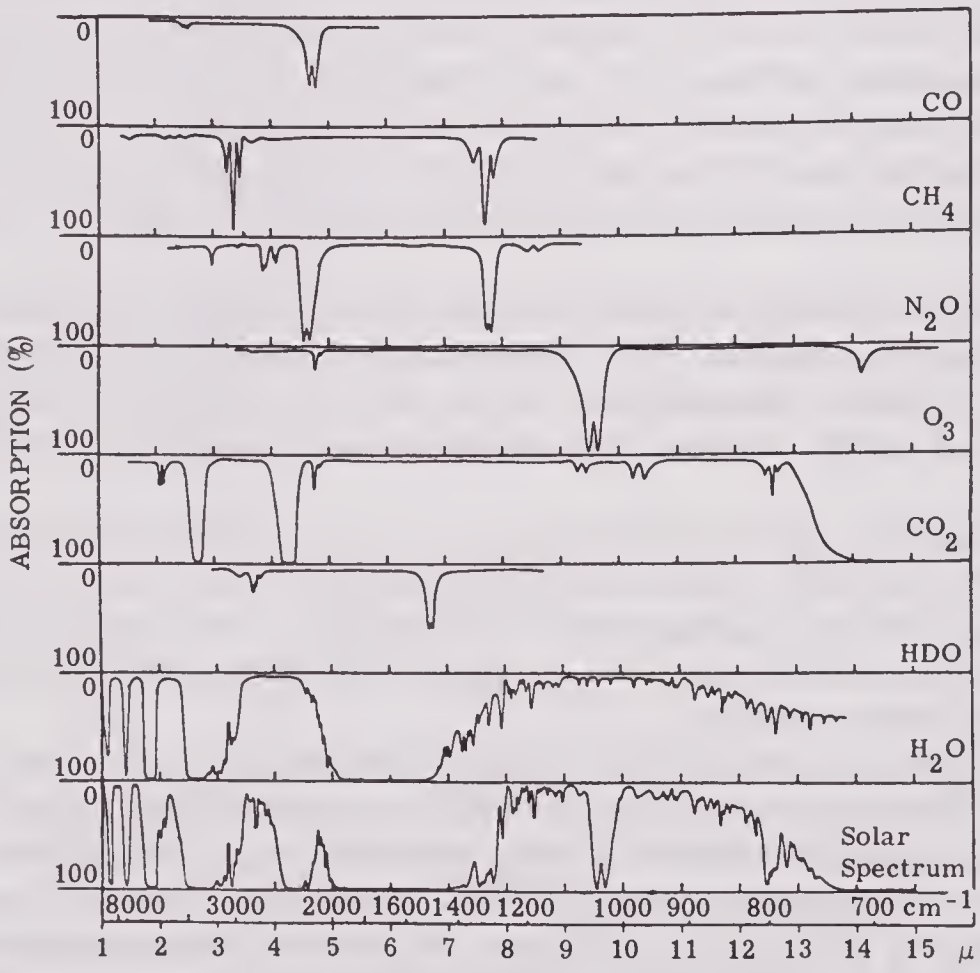


FIG. 6-59. Low-resolution solar spectrum from 1 to 24 μ.

From 3.50 to 3.85 μ (Fig. 6-62), many strong lines of the ν_1 fundamental band of HDO are present, in additio? to some CH_4 absorption ($\nu_2 + \nu_4$ at 3.55 μ) and a weak N_2O combination band ($\nu_2 + \nu_3$ at 3.57 μ). The Q-branch of the HDO band appears as a weak cluster of lines near 3.67 μ. All lines marked × in Fig. 6-62 are due to HDO.

The region from 3.85 to 4.20 μ (Fig. 6-63) contain $2\nu_1$ of N_2O at 3.91 μ and $\nu_1 + 2\nu_2$ of N_2O at 4.06 μ. It also contains the ν_3 fundamental of CO_2, beginning at 4.18 μ.

Virtually complete absorption by CO_2 occurs in the region from 4.19 μ (Fig. 6-63) to 4.45 μ (Fig. 6-64). From 4.43 to 4.48 μ are some high-J lines in the P-branch of ν_3 of $^{13}CO_2$, centered at 4.38 μ. The intense ν_3 fundamental band of N_2O is at 4.49 μ, with a weaker band $(\nu_3 + \nu_2) - \nu_2$ of N_2O centered near 4.52 μ.

Near 4.66 μ (Fig. 6-64) is the fundamental band of CO. Circles above the lines in (Fig. 6-64) indicate CO transitions of solar origin. The absorption at approximately 4.7 μ is due to ozone. The ozone absorption lines are indicated by circles below the spectrum. H_2O absorption lines between about 4.64 and 4.68 μ are indicated by H.

A weak CO_2 band is present near 4.8 μ (Fig. 6-65) and a very weak CO_2 band near 5.2 μ (Fig. 6-66). Beyond about 5.2 μ, there is strong absorption from the edge of the ν_2 fundamental band of H_2O which is centered at 6.2 μ. Absorption by water vapor is complete between 5.5 and 6.9 μ. From 6.9 to 7.65 μ (Fig. 6-67), the primary absorption is the edge of the 6.2-μ H_2O band, although absorption due to the ν_4 vibrational-rotation band of CH_4, centered at 7.65 μ, is present. The region between 7.65 μ (Fig. 6-68) and 9.0 μ (Fig. 6-69) possesses relatively high transmission, except for the overtone band $2\nu_2$ of N_2O centered at 8.56 μ.

Figure 6-70 shows very intense absorption due to the ν_3 band of O_3 centered at 9.60 μ. The very faint structure observed between 8.90 μ (Fig. 6-69) and 9.15 μ (Fig. 6-70) is due to the very weak ν_1 band of O_3.

The $\nu_3 - 2\nu_2$ difference band of CO_2, centered at 9.4 μ, is also shown in Fig. 6-70. In a difference band, the individual absorption line is not caused by a quantum transition from the ground state of the molecule to an excited level but rather to a transition from an excited level to a higher excited level. Because the intensity of an absorption line depends very strongly on the population of the energy level from which the transition originates, and since this population decreases as temperature decreases, difference bands fade out very rapidly as temperature decreases. Thus, a band such as the 9.40-μ CO_2 band, although it may cause significant absorption near ground level where the ambient temperature may be 300°K, may have very little intensity near 100,000 ft where the temperature is about 200°K.

The region from 9.75 to 10.6 μ (Fig. 6-71) shows more of the structure of the 9.6-μ ozone band. It also shows another difference band of CO_2, $\nu_3 - \nu_1$, centered at 10.4 μ.

Relatively high transmission is present in the region from 12.2 to 13.25μ (Fig. 6-72 and 6-73). Figure 6-73 shows a CO_2 difference band, $\nu_1 - \nu_2$, centered near 13.2 μ and a much weaker CO_2 difference band, $(\nu_1 + \nu_2) - 2\nu_2$, centered near 12.6 μ.

Figure 6-74 shows more of the structure of the $\nu_1 - \nu_2$ difference band of CO_2 and, starting near 13.7 μ, the absorption by the very intense ν_2 fundamental band of CO_2. Absorption by this band is complete to about 17 μ.

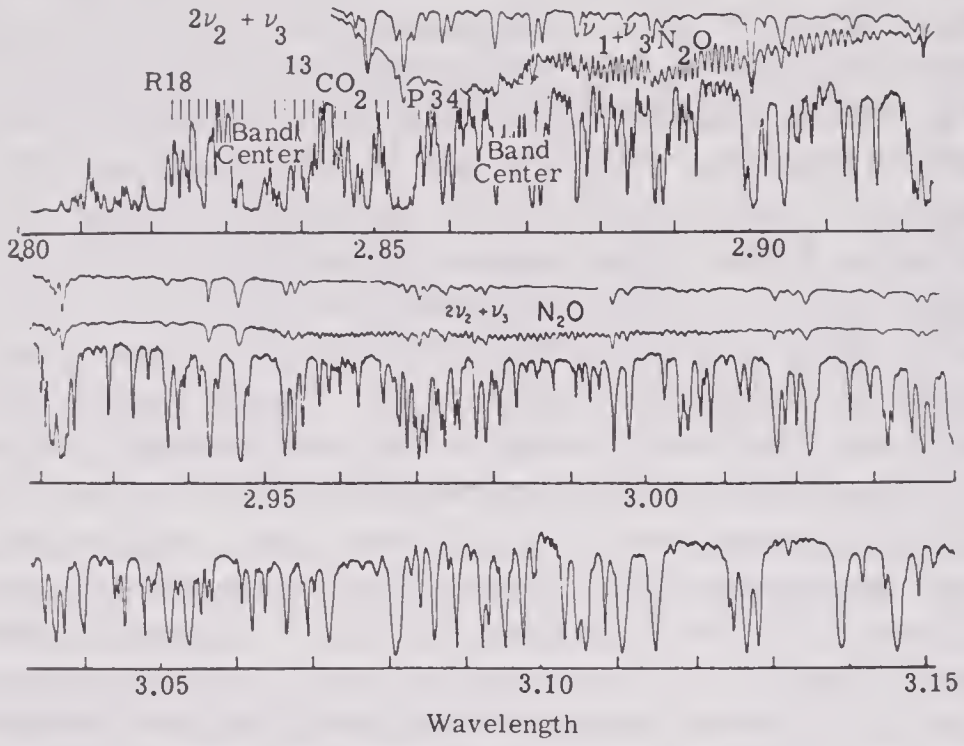


FIG. 6-60. Solar spectrum from 2.80 to 3.15 μ (lowest curve); laboratory spectrum of H_2O (top curve); laboratory spectrum of N_2O (middle curve).

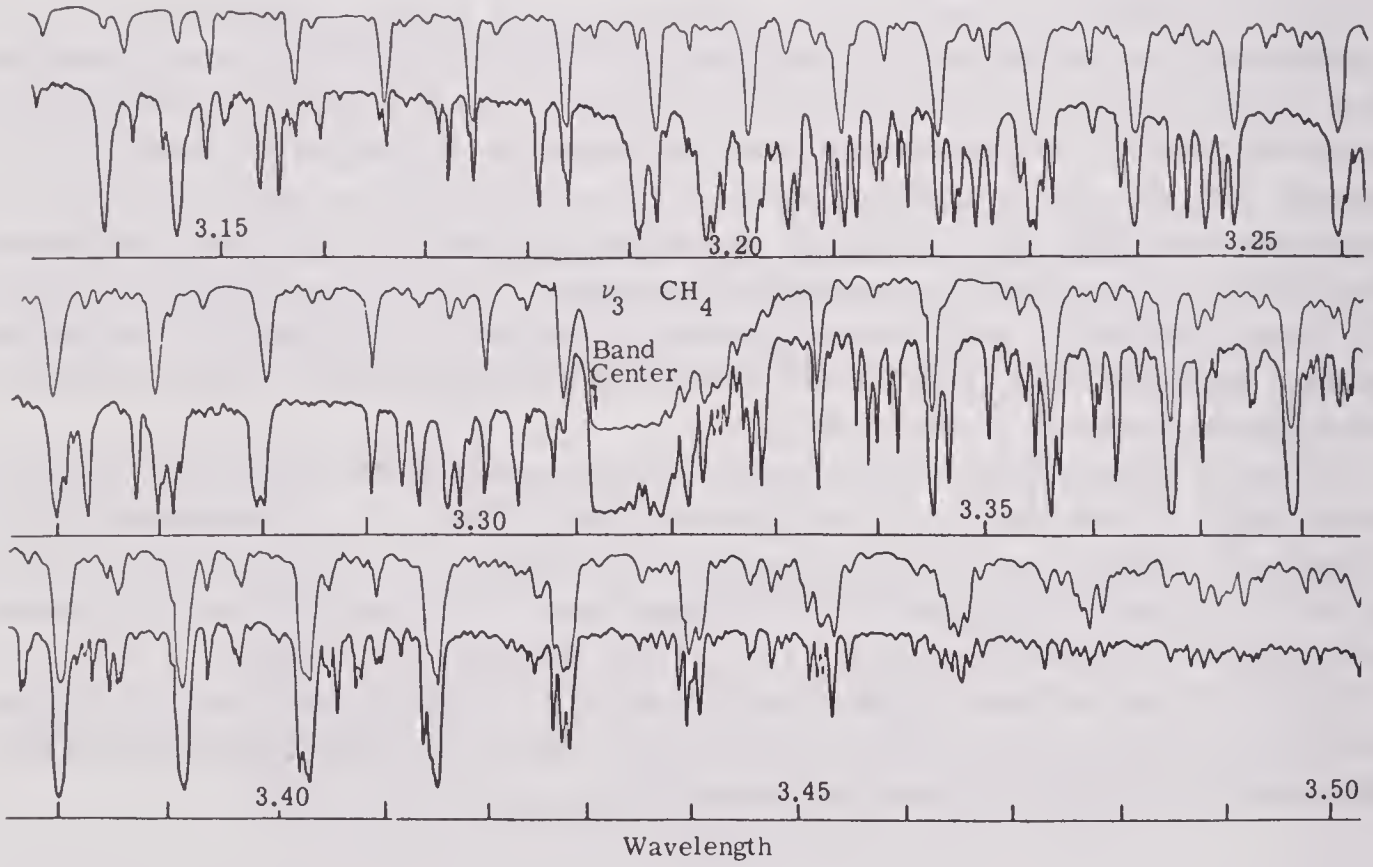


FIG. 6-61. Solar spectrum from 3.15 to 3.50 μ (lower curve); laboratory spectrum of CH_4 (upper curve).

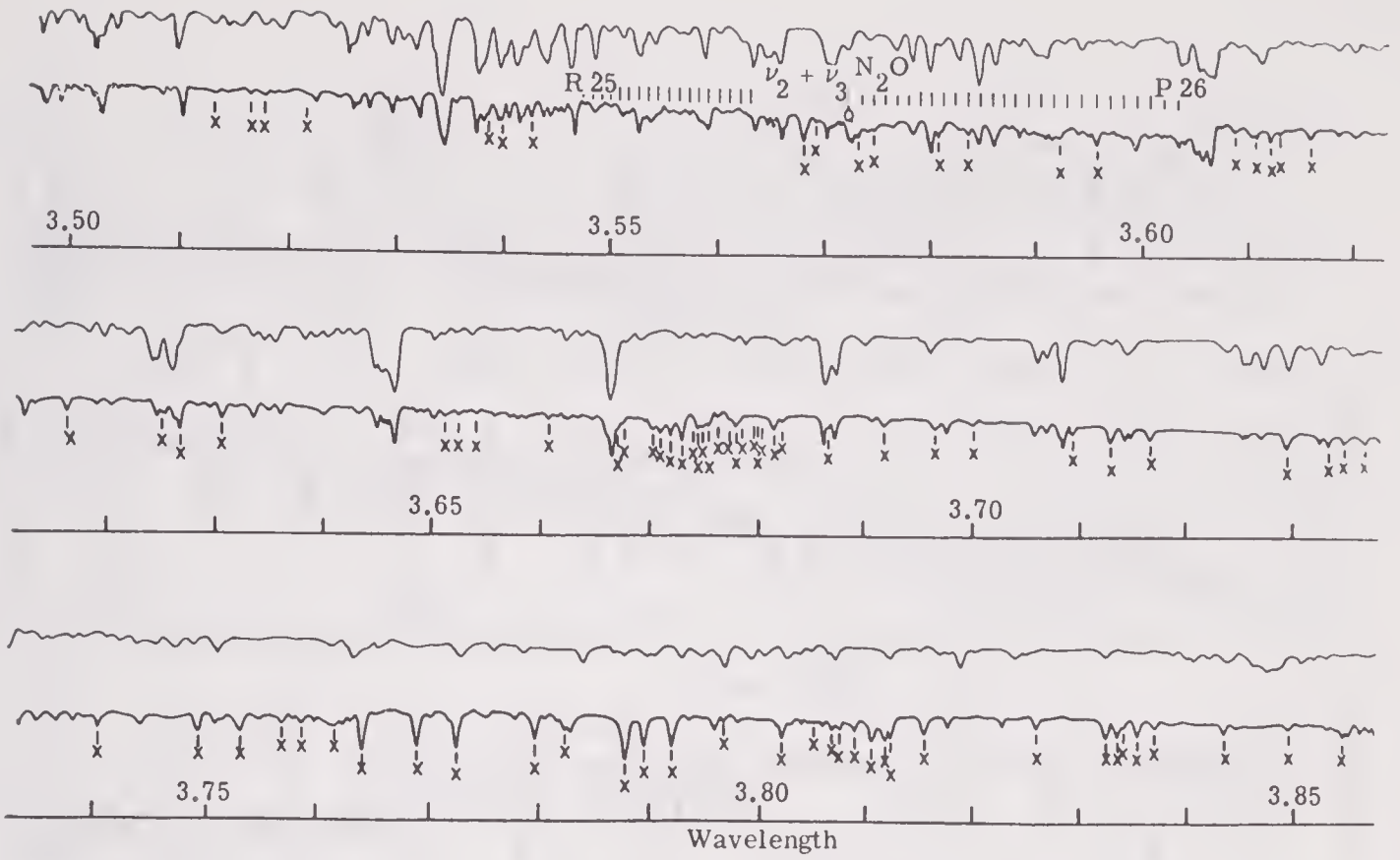


FIG. 6-62. Solar spectrum 3.50 to 3.85 μ (lower curve); laboratory spectrum of CH_4 (lower curve).

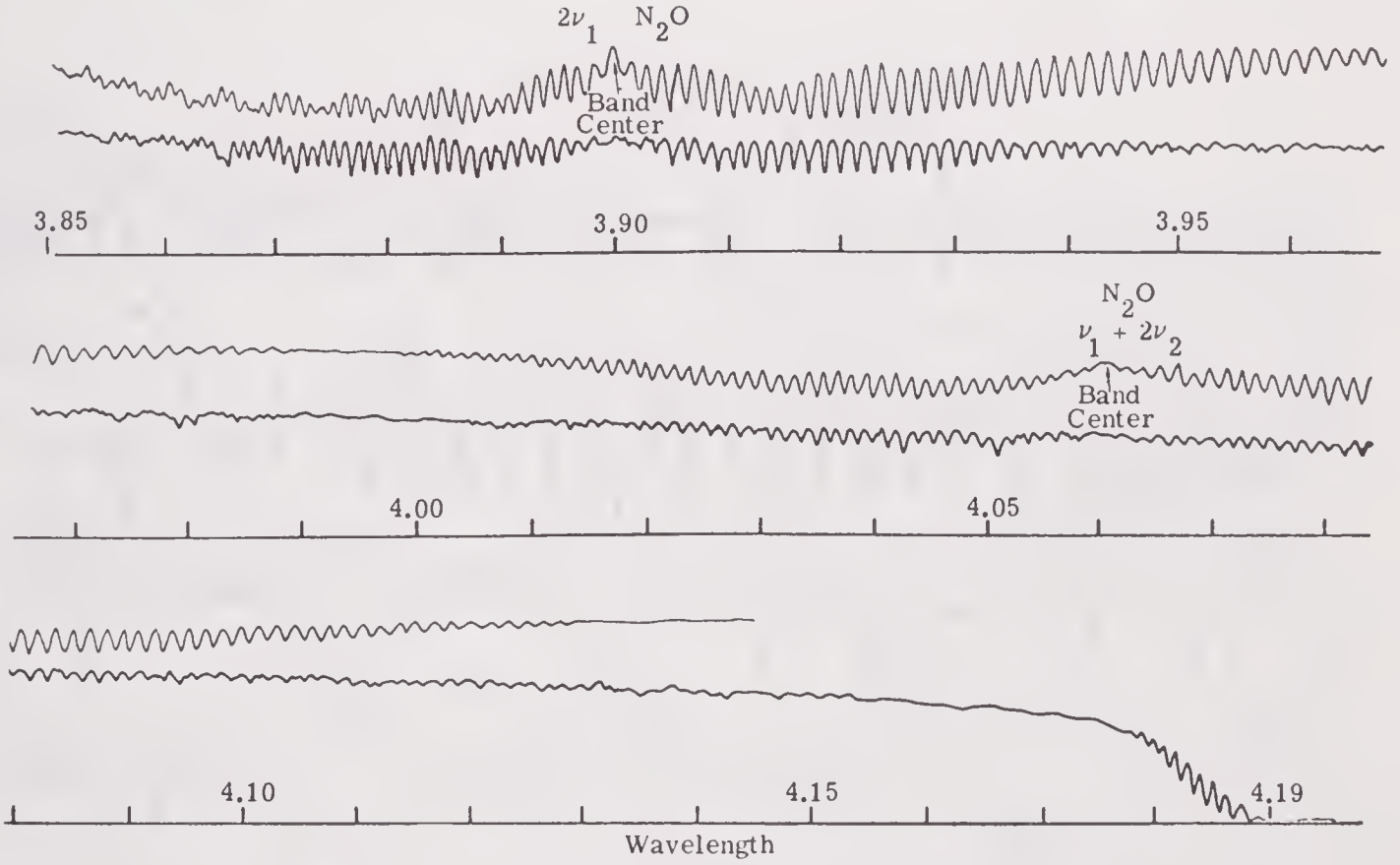


FIG. 6-63. Solar spectrum from 3.85 to 4.19 μ (lower curve); laboratory spectrum of N_2O (upper curve).

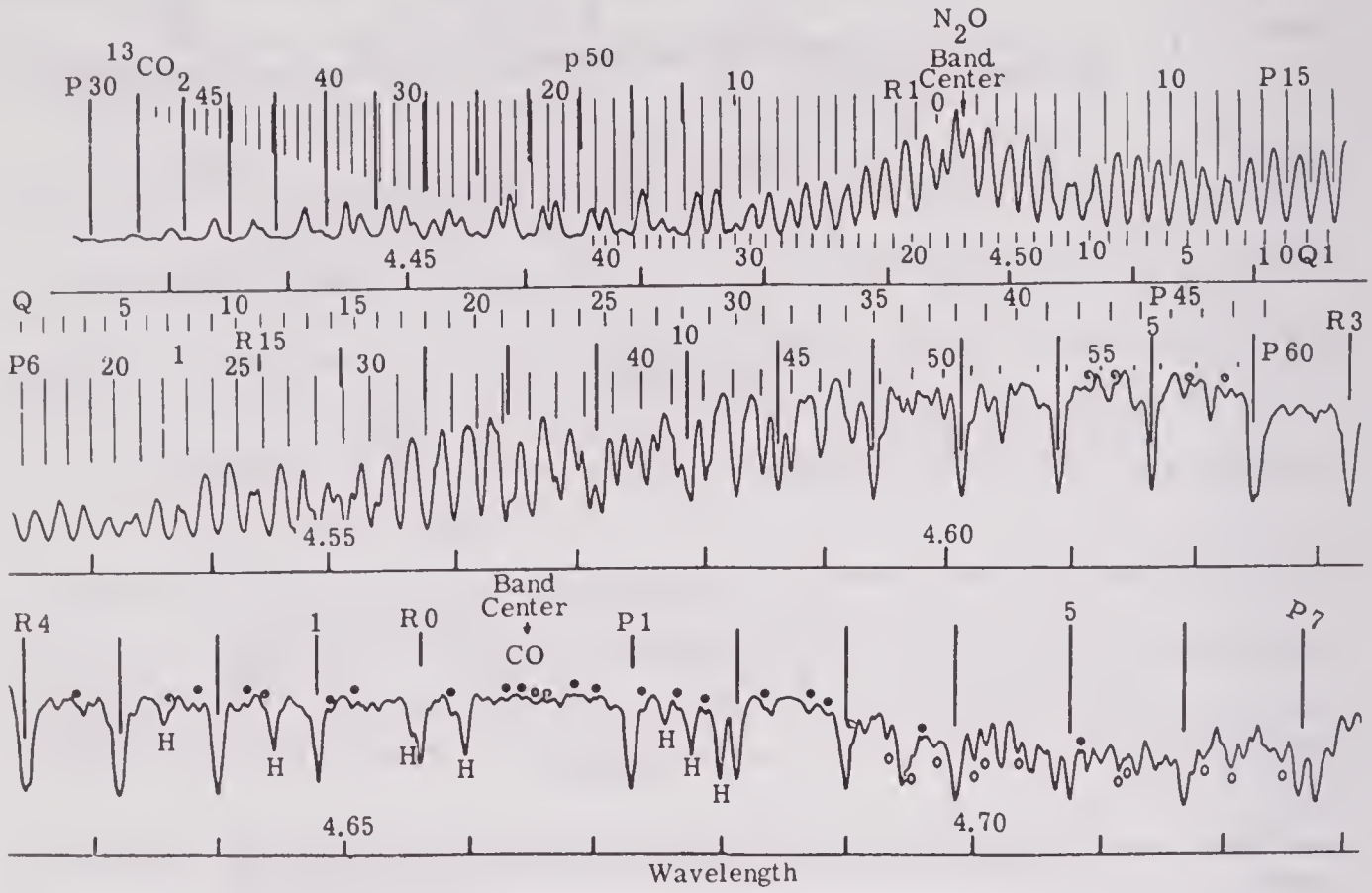


FIG. 6-64. Solar spectrum from 4.43 to 4.73 μ. Absorption structure by CO_2, N_2O, and CO_2 are shown schematically.

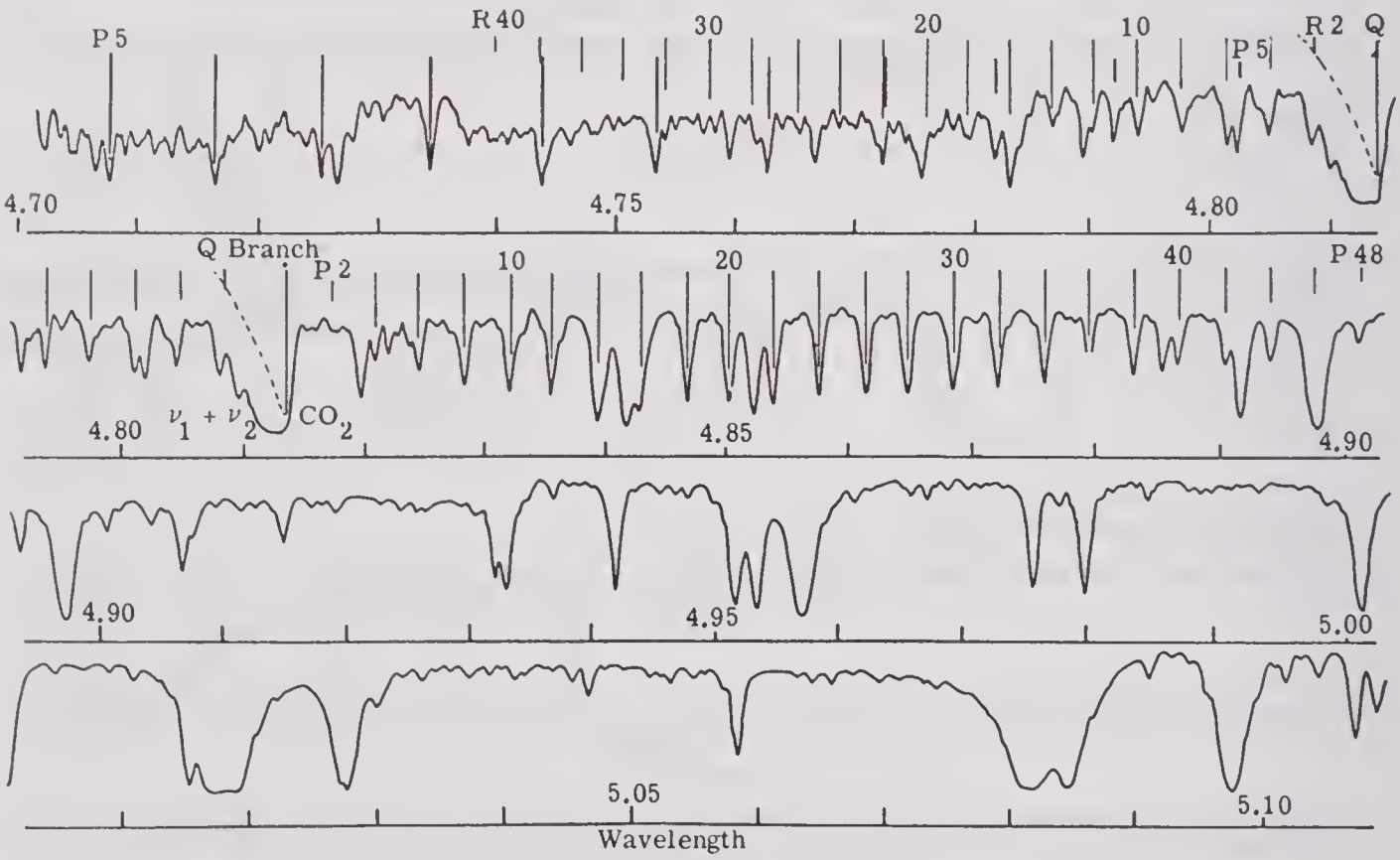


FIG. 6-65. The solar spectrum from 4.70 to 5.11 μ.

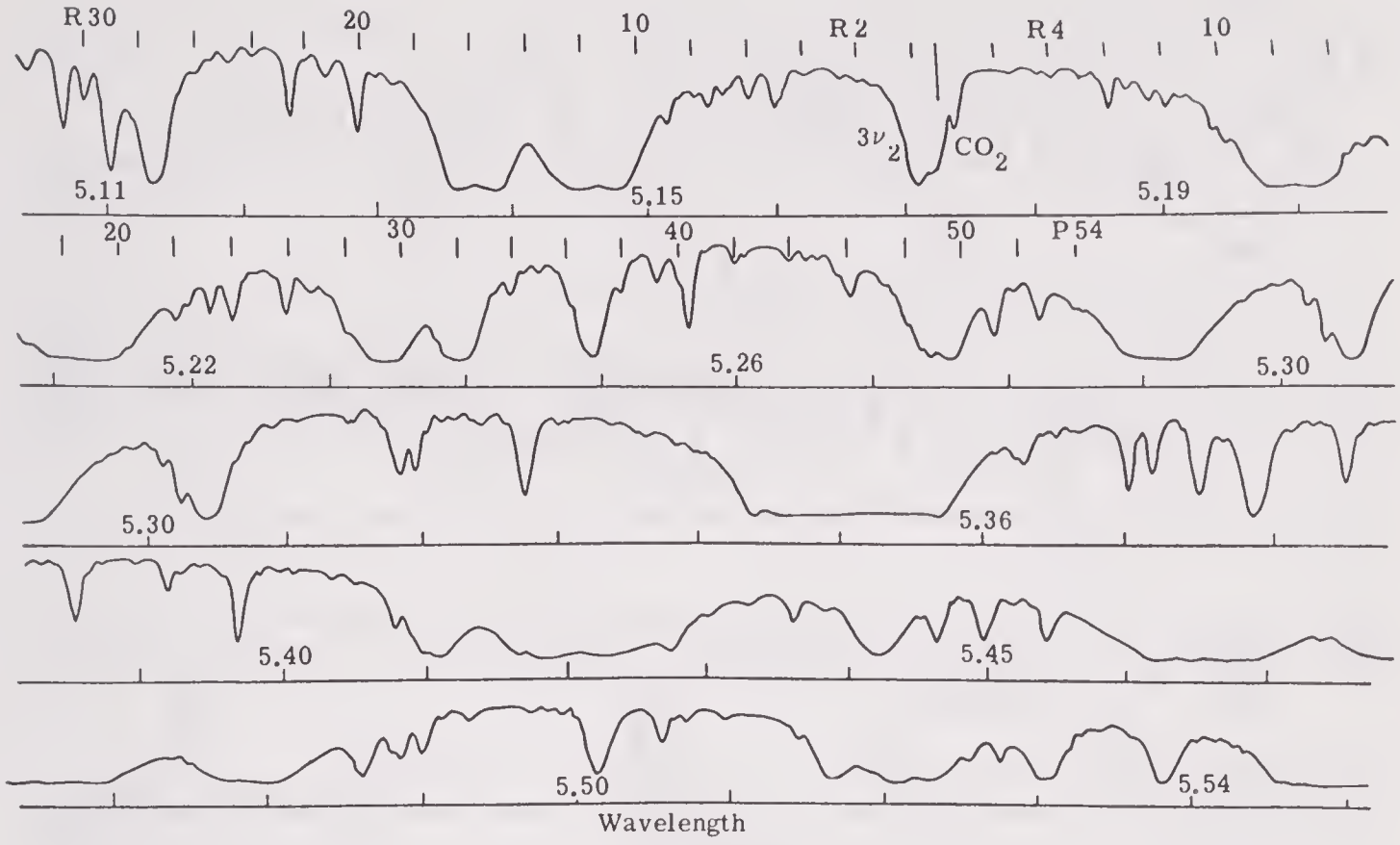


FIG. 6-66. The solar spectrum from 5.11 to 5.55 μ.

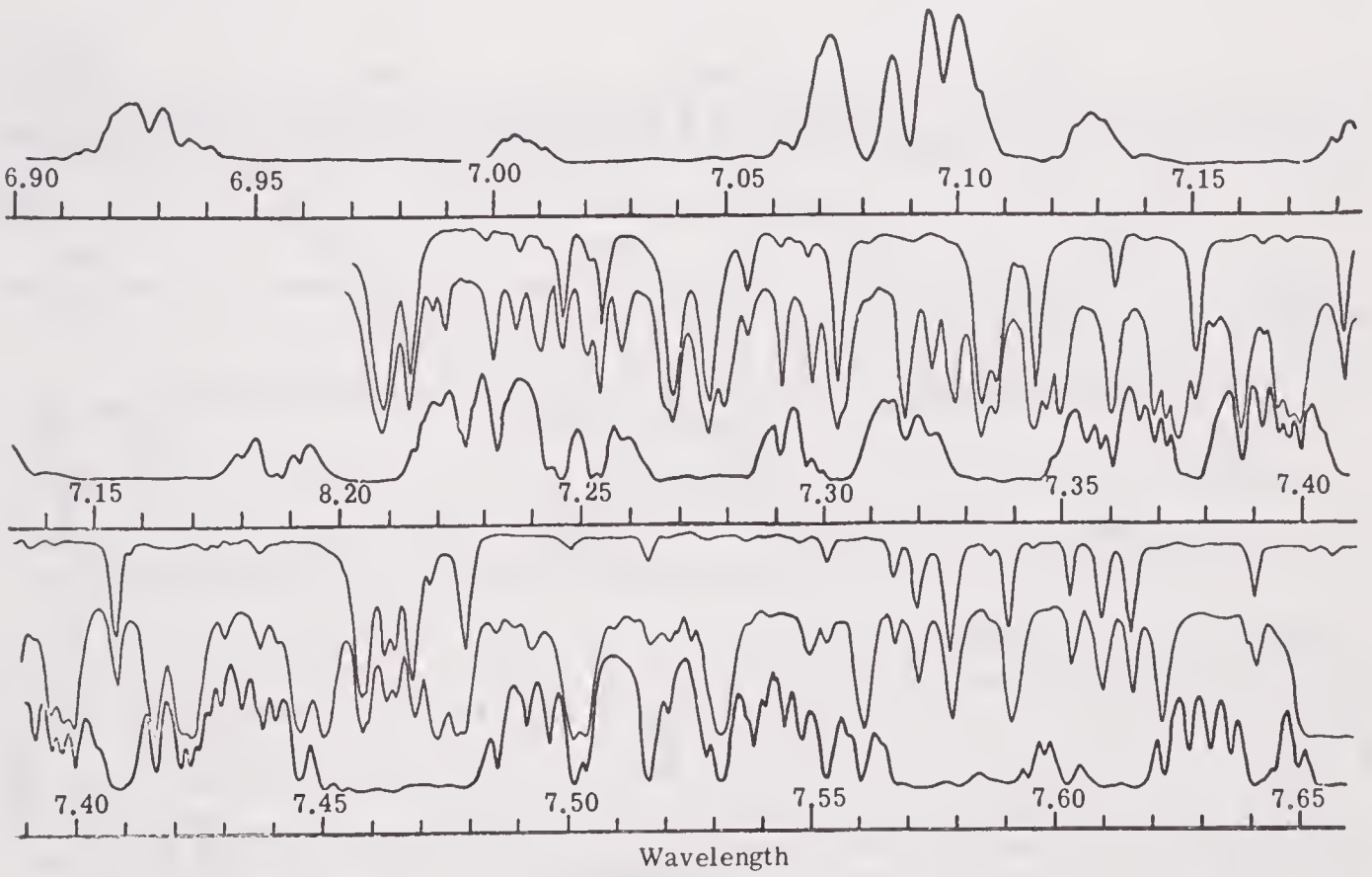


FIG. 6-67. Solar spectrum from 6.90 to 7.65 μ (lowest curve); laboratory spectrum of H_2O (top curve); laboratory spectra of CH_4 (middle curve).

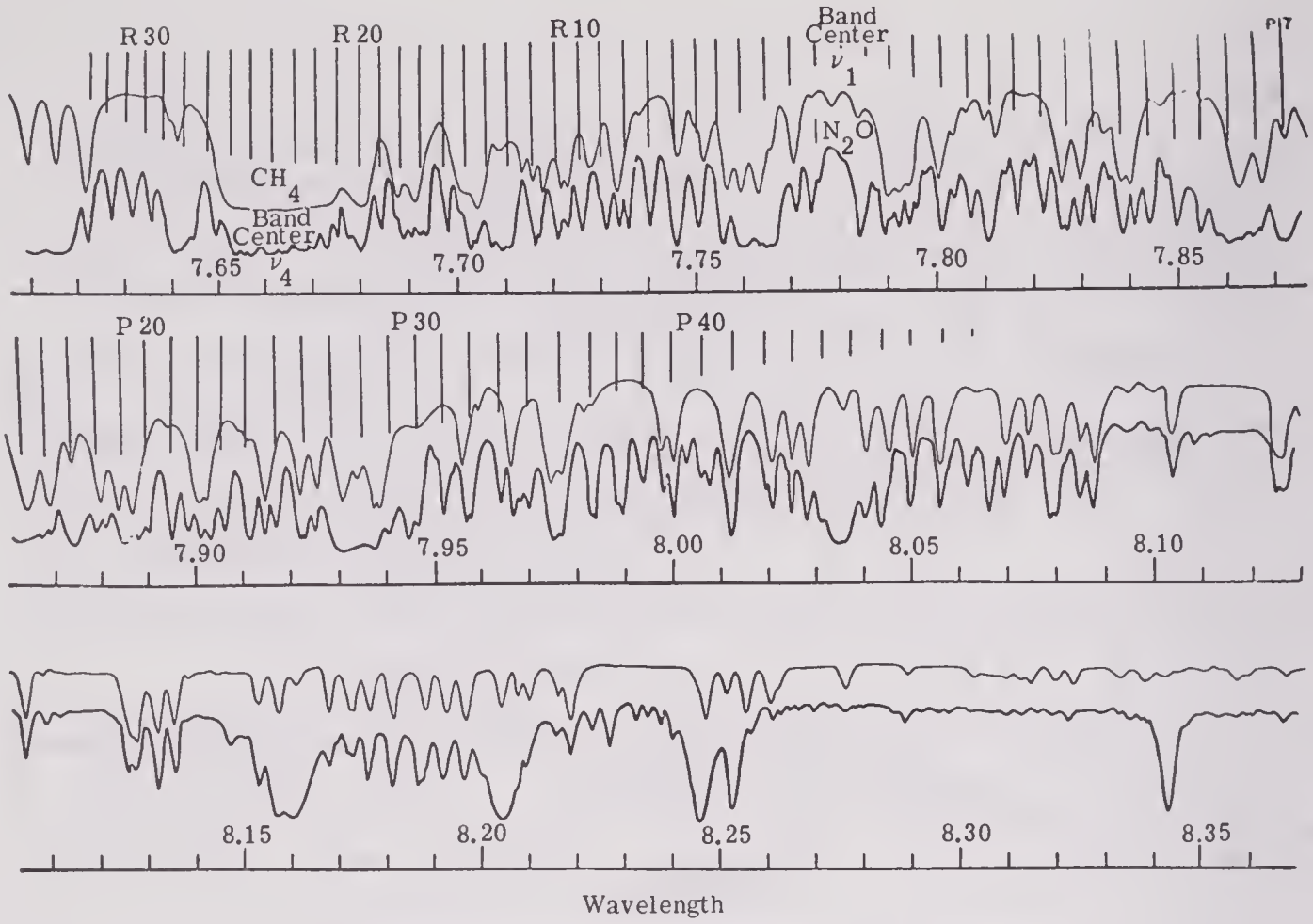


FIG. 6-68. Solar spectrum from 7.65 to 8.35 μ (lower curve); laboratory spectrum of CH_4 (upper curve); absorption structure of N_2O indicated schematically.

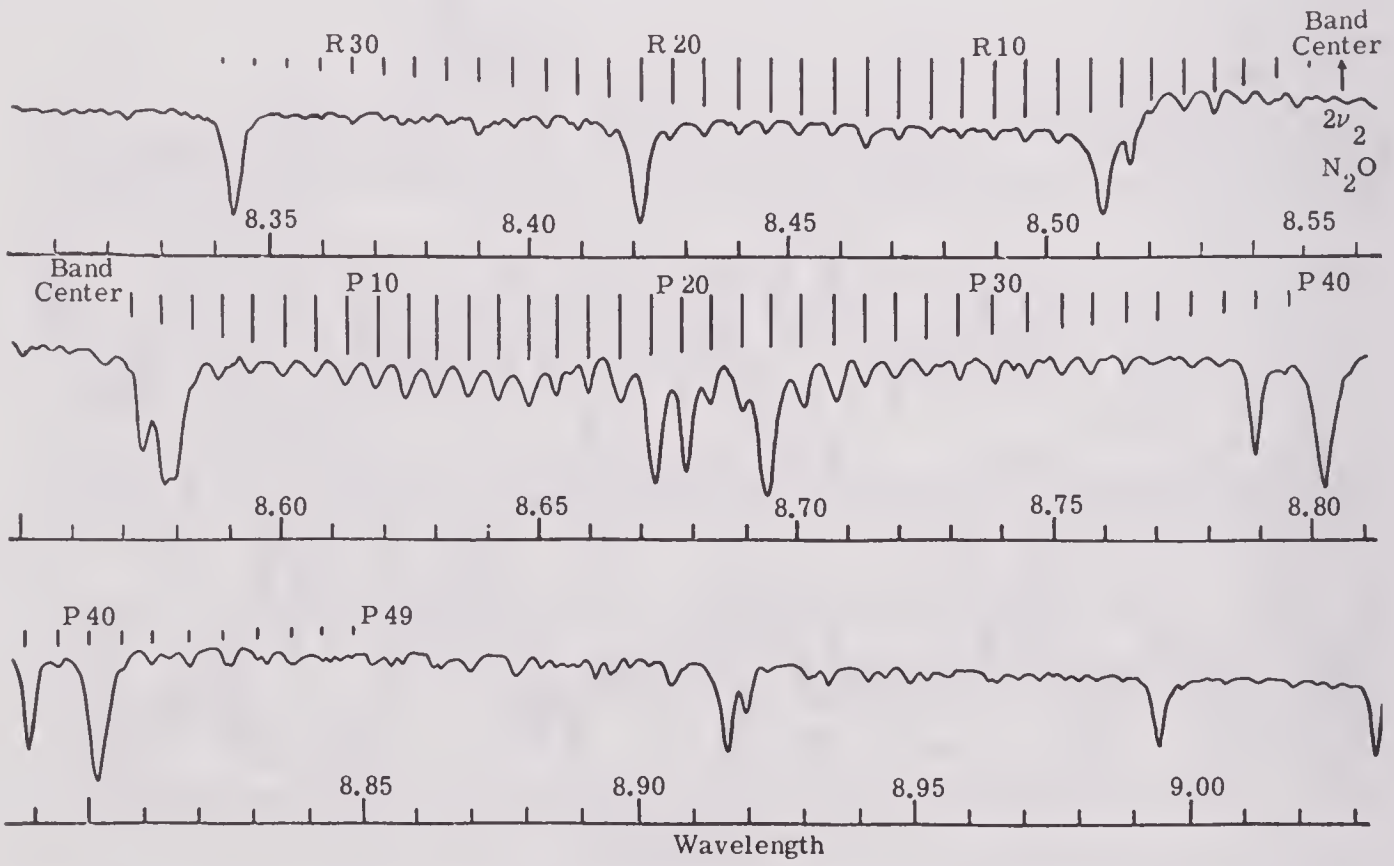


FIG. 6-69. Solar spectrum from 8.35 to 9.03 μ. Absorption structure of N_2O indicated schematically.

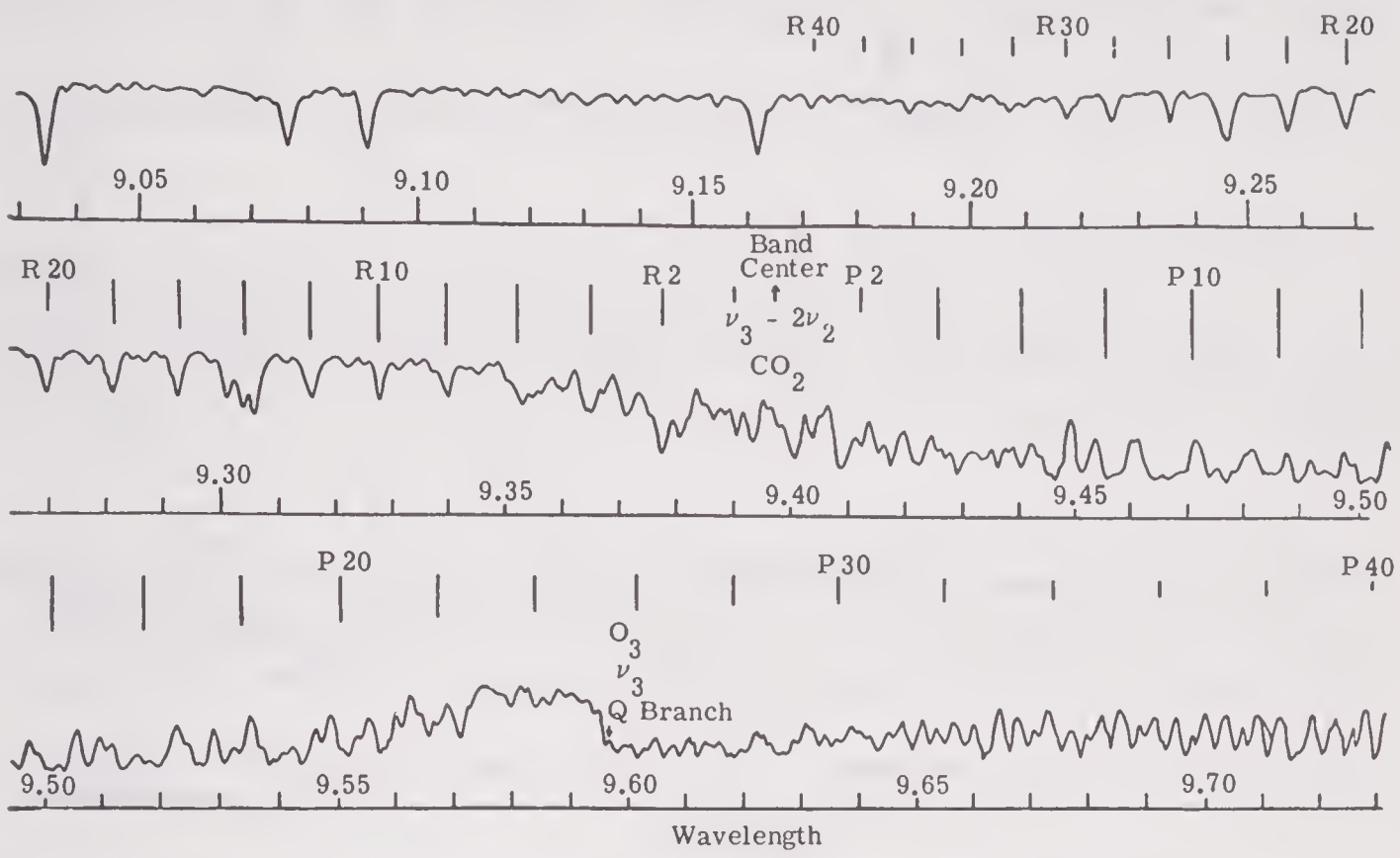


FIG. 6-70. Solar spectrum from 9.03 to 9.73 μ. Absorption structure of CO_2 indicated schematically.

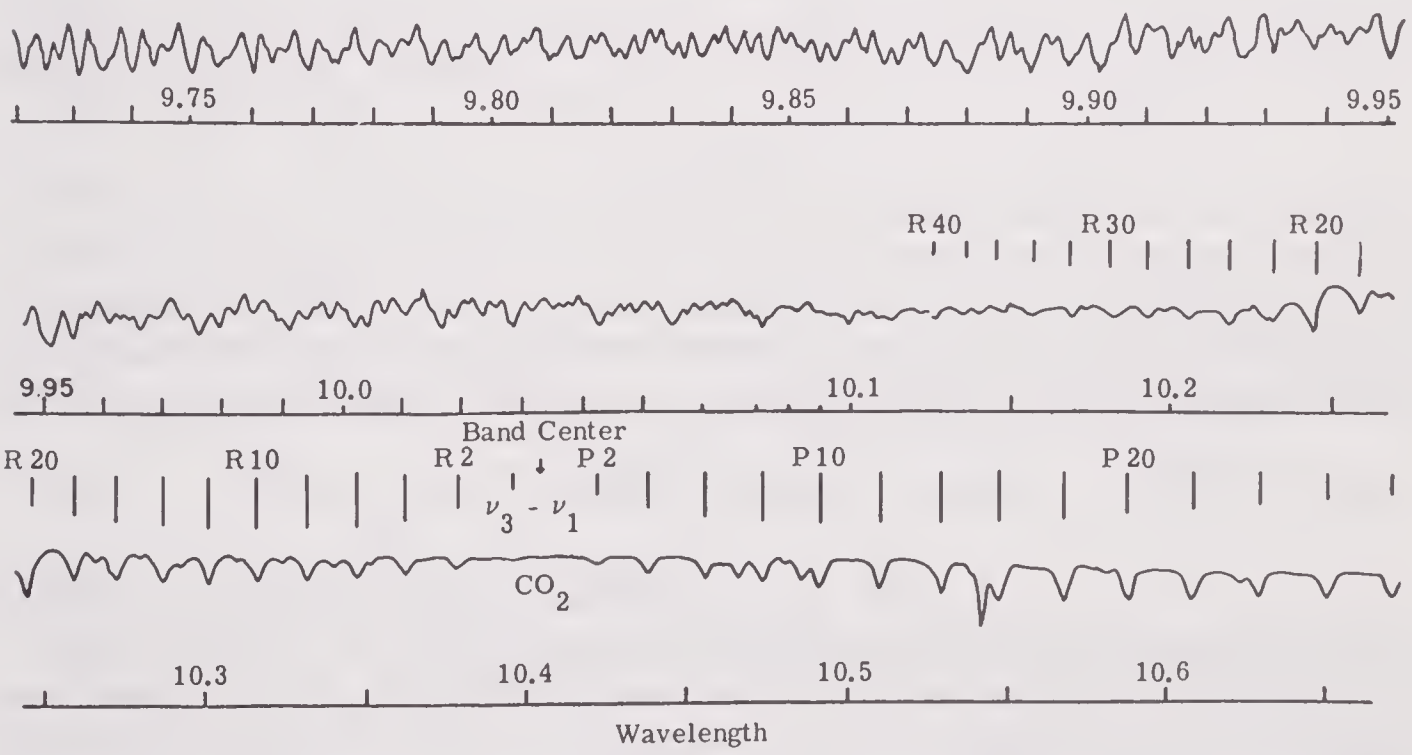


FIG. 6-71. Solar spectrum from 9.72 to 10.7 μ. Absorption spectra of CO_2 indicated schematically.

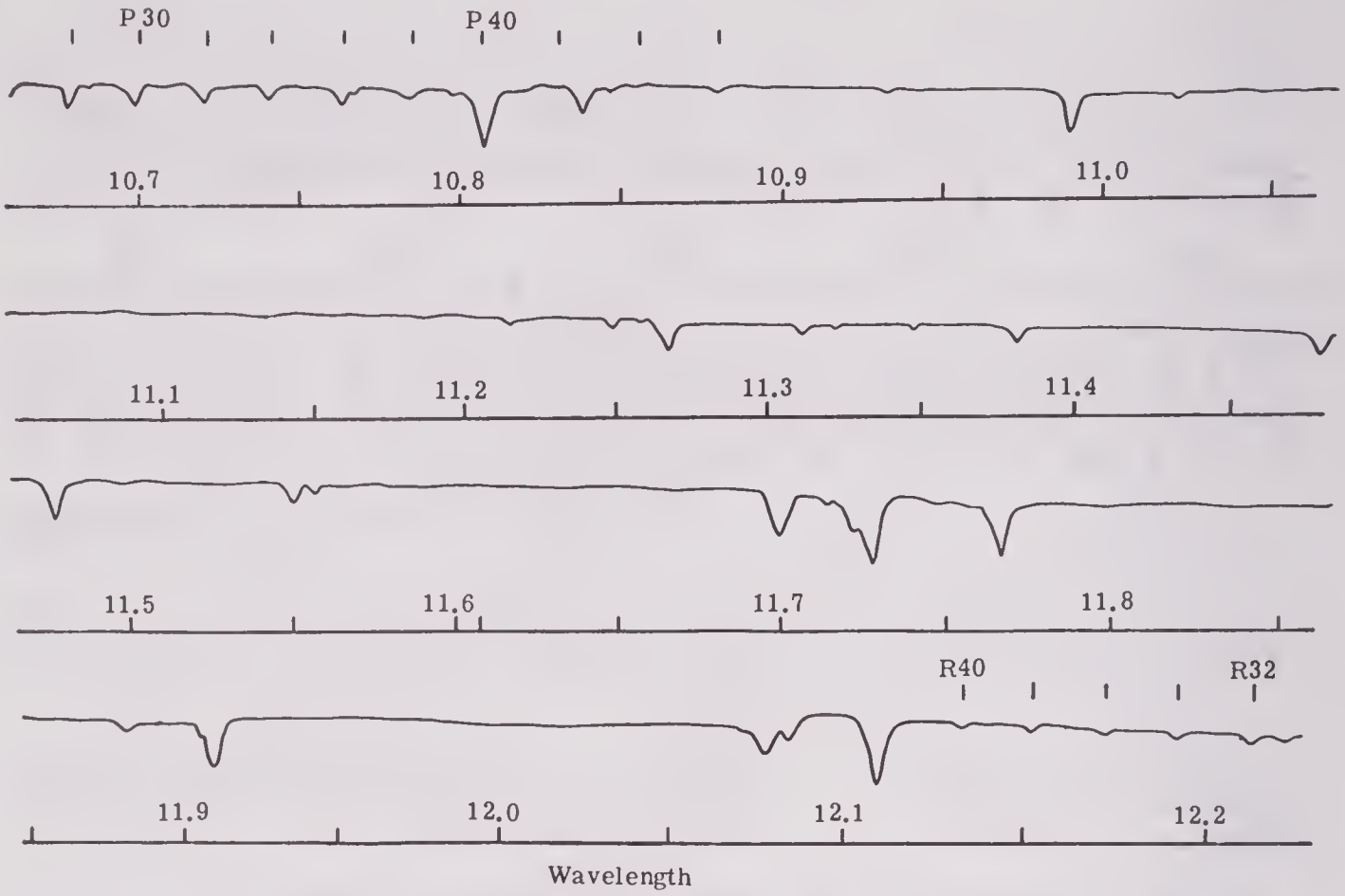


FIG. 6-72. Solar spectrum from 10.7 to 12.2 μ. Absorption structure of CO_2 indicated schematically.

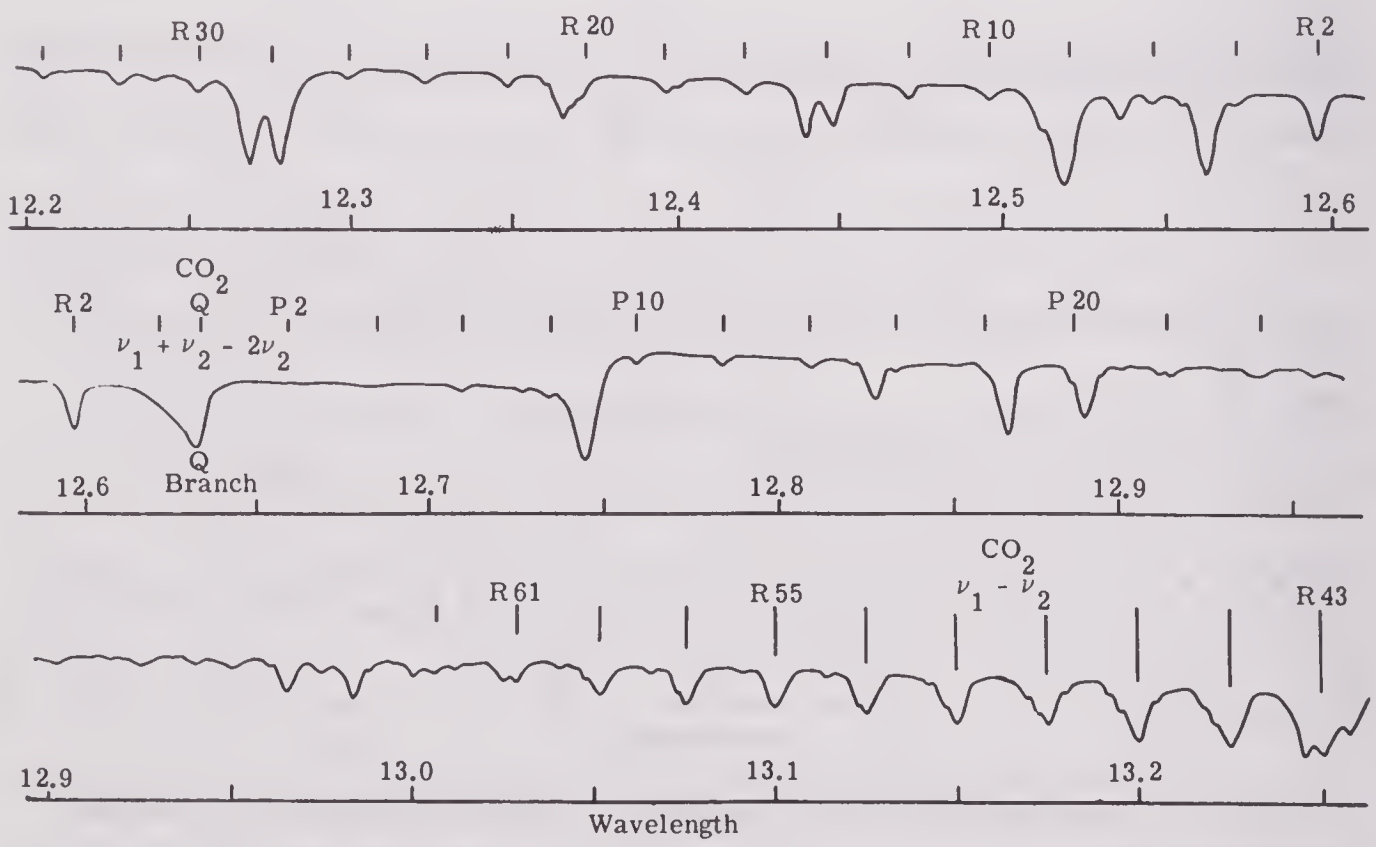


FIG. 6-73. Solar spectrum from 12.2 to 13.3 μ. Absorption structure for CO_2 indicated schematically.

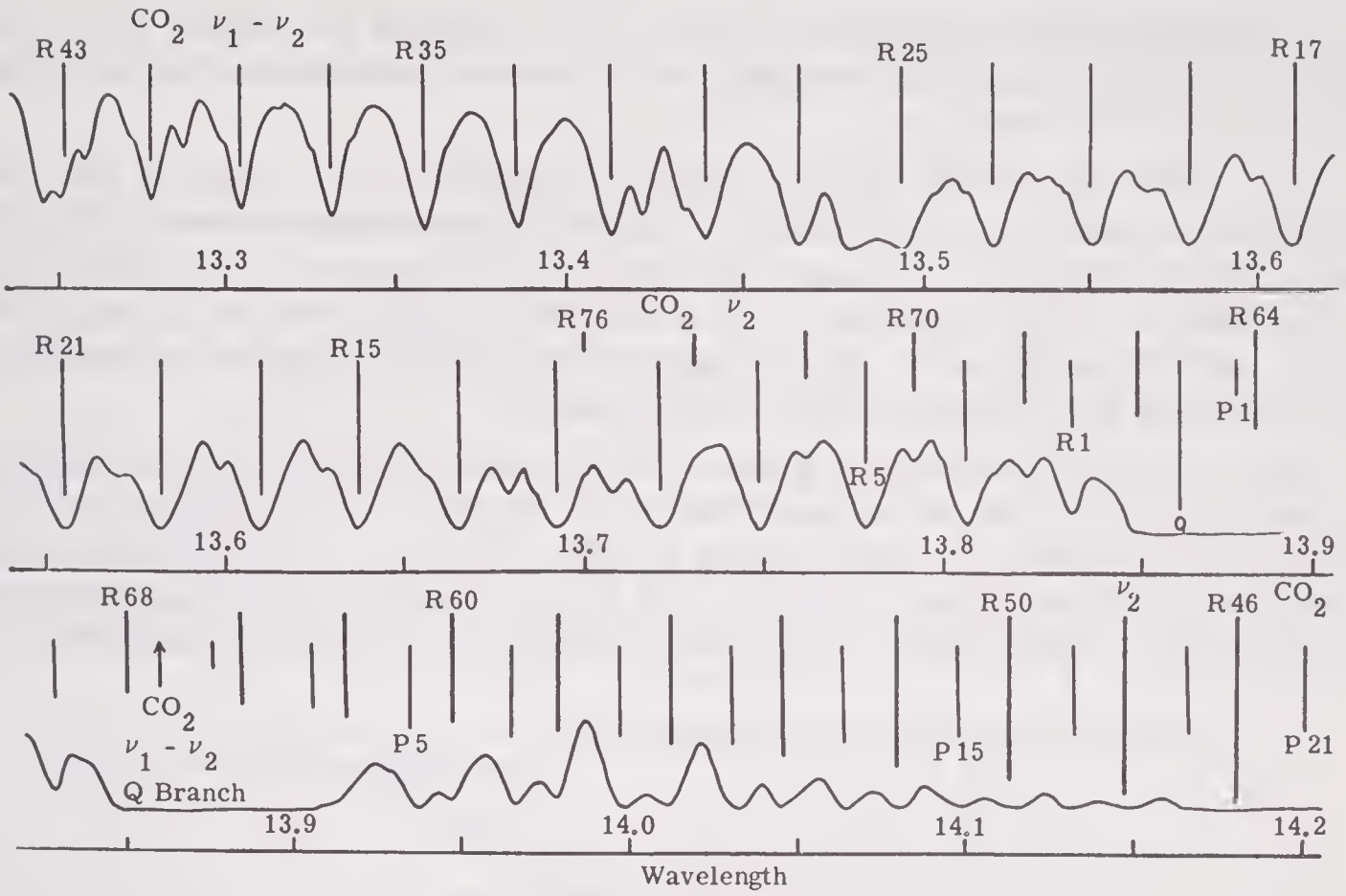


FIG. 6-74. Solar spectrum from 13.25 to 14.2 μ. Absorption structure of CO_2 indicated schematically.

6.8. Total Absorption (Laboratory Measurements) [104-107]

Total absorption data provide a means to predict absorption for known paths through known absorbing gases and to test the validity of theories describing absorption phenomena. The total absorption of an absorption band is the area under the curve obtained when the fractional absorption at a given frequency is plotted against frequency. Usually the integral

$$\int_{\nu_1}^{\nu_2} A(\nu)\, d\nu$$

is called the total absorption of an absorption band, defined by the limits ν_1 and ν_2, and is expressed in frequency units. Sometimes, however, the integral is referred to as the equivalent bandwidth of the absorption because the same integral can be considered as applying to an equivalent band having complete absorption over a frequency interval

$$\Delta\nu = \int_{\nu_1}^{\nu_2} A(\nu)\, d\nu$$

The former definition is used in this section.

Data presented in this section show the functional relationship between total absorption $\int A(\nu)\, d\nu$ and absorber concentration w, partial pressure of the absorbing gas p, total pressure P, which includes the partial pressure of absorbing and nonabsorbing gases, and the absolute temperature of the gas T. This relationship is expressed as

$$\int_{\nu_1}^{\nu_2} A(\nu)\, d\nu = \phi(w, p, P, T)$$

and is described for various wavelength regions of high characteristic absorption.

The data were measured using multiple-traversal cells containing absorbing gas the partial pressure of which could be varied. High-altitude conditions were simulated by

proper vacua by adding broadening gases such as nitrogen and oxygen to the cell. Path lengths from 1.5 to 48 m were achieved by successive reflections of radiation back and forth through the cell.

6.8.1. Total Absorption by CO_2. Strong absorption by CO_2 exists in the 2.7-μ (3660-cm^{-1}) region, the 4.3-μ (2350-cm^{-1}) region, and the region between 11.4 μ (875 cm^{-1}) and 20 μ (495 cm^{-1}). Weaker absorption bands are present at 1.4 μ (6975 cm^{-1}), 1.6 μ (6230 cm^{-1}), 2.0 μ (4983 cm^{-1}), 4.8 μ (2075 cm^{-1}), 5.2 μ (1930 cm^{-1}), 9.4 μ (1064 cm^{-1}), and 10.4 μ (961 cm^{-1}). In the figures that follow, P_e = equivalent pressure as used by Burch and coworkers, and P is total pressure.

6.8.1.1. *The* 2.7-μ (3660-cm^{-1}) *Region.* The CO_2 absorption in the 2.7-μ region is caused primarily by two strong absorption bands, the $2\nu_2 + \nu_3$ band centered at 2.77 μ (3609 cm^{-1}) and the $\nu_1 + \nu_3$ band centered at 2.69 μ (3716 cm^{-1}). Total absorption for the 2.77-μ (3609-cm^{-1}) band is shown in Fig. 6-75, and total absorption for the 2.69-μ band (3716 cm^{-1}) band in Fig. 6-76. Total absorption for the entire 2.7-μ (3660 cm^{-1})

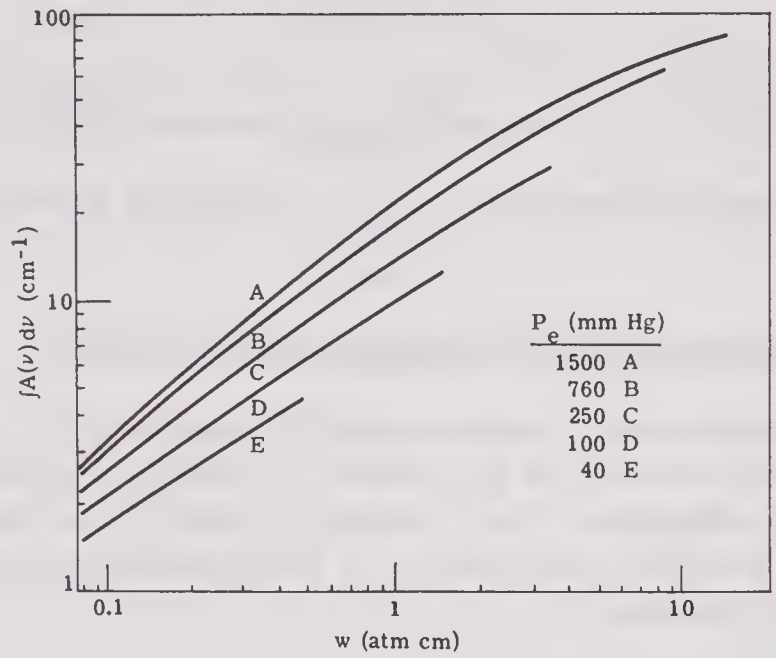


FIG. 6-75. Total absorption *vs.* absorber concentration of the 2.77-μ (3609-cm^{-1}) CO_2 band.

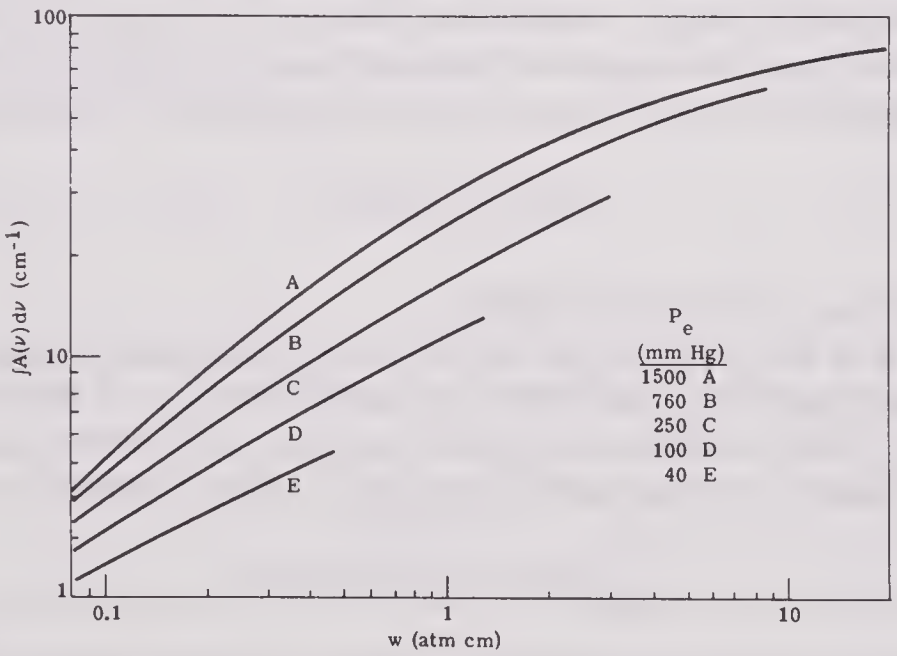


FIG. 6-76. Total absorption *vs.* absorber concentration for the 2.69-μ (3716-cm^{-1}) CO_2 band.

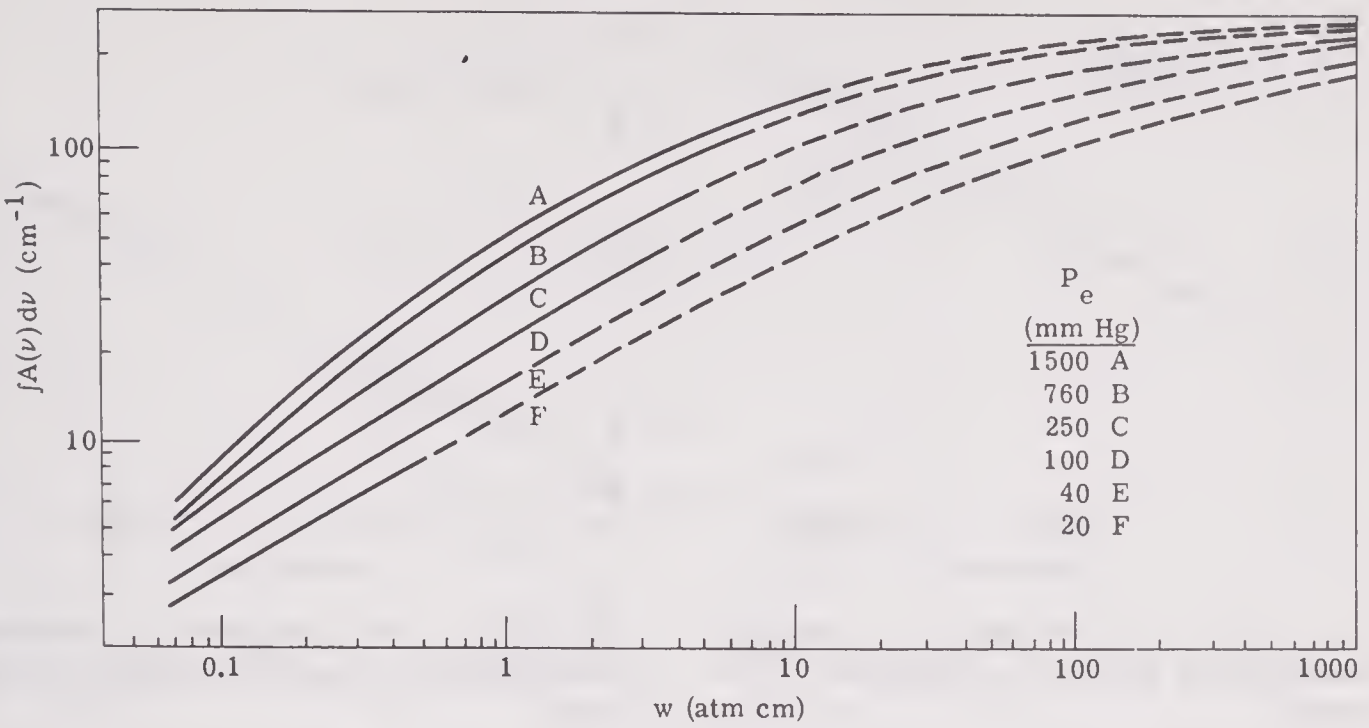


FIG. 6-77. Total absorption *vs.* absorber concentration for the 2.7-μ (3660-cm^{-1}) CO_2 region (2.77-μ band + 2.69-μ band).

region, *i.e.*, the 2.77-μ (3609-cm^{-1}) band plus the 2.69-μ (3716-cm^{-1}) band, is shown in Fig. 6-77. For a total absorption of more than 3 cm^{-1}, the curves in all of the illustrations are estimated to be accurate within $\pm$5%; between 3 and 10 cm^{-1}, the estimated accuracy is $\pm$10%. For total absorption values less than 3 cm^{-1}, the estimated accuracy is $\pm$20%.

6.8.1.2. *The* 4.3-μ (2350-cm^{-1}), 4.8-μ (2075-cm^{-1}), *and* 5.2-μ (1930-cm^{-1}) *Bands.* The 4.3-μ (2350-cm^{-1}) CO_2 band causes almost complete absorption between about 4.19 μ (2386 cm^{-1}) and 4.45 μ (2250 cm^{-1}). Total absorption curves, which are estimated to be accurate within $\pm$5% above 10 cm^{-1} total absorption, $\pm$10% between 3 and 10 cm^{-1} total absorption, and 20% below 3 cm^{-1} total absorption, are shown in Fig. 6-78.

Total absorption curves for the 4.8-μ (2075-cm^{-1}) and the 5.2-μ (1930-cm^{-1}) band are shown in Fig. 6-79 and 6-80, respectively. These bands are very weak and are significant only for high values of absorber concentration.

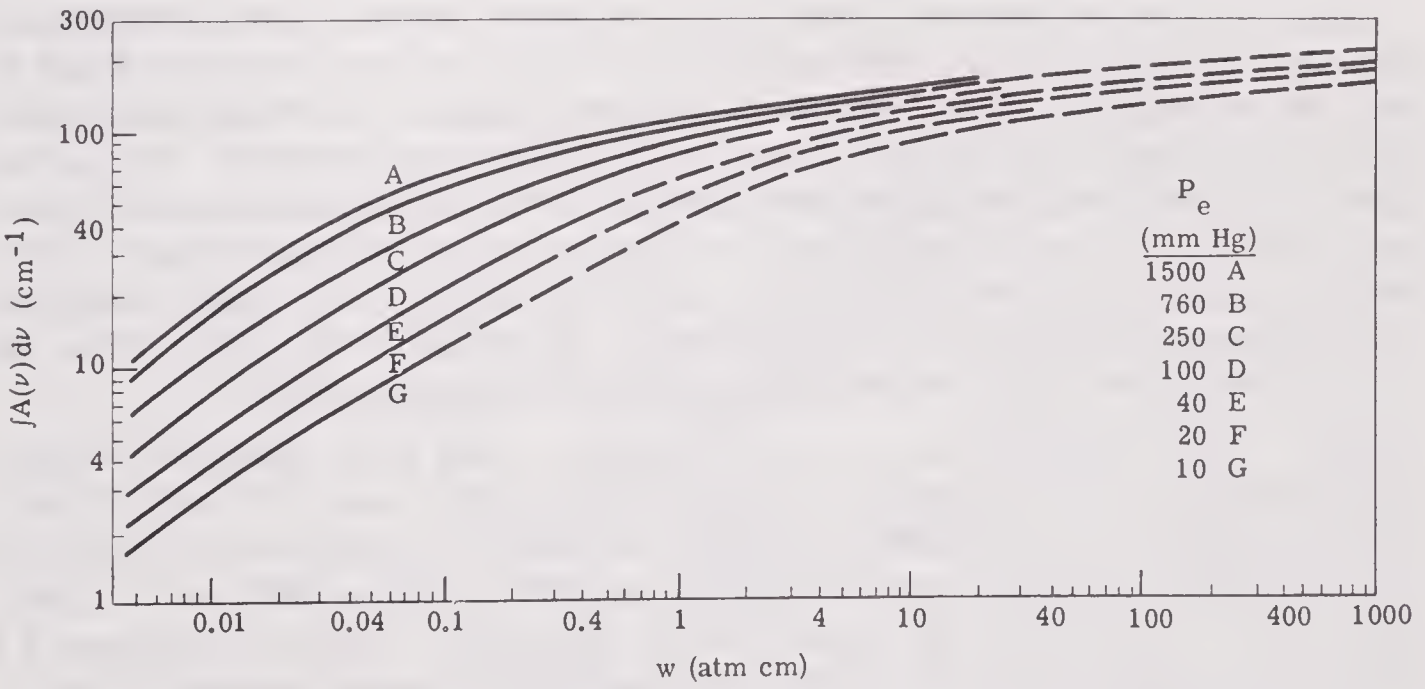


FIG. 6-78. Total absorption *vs.* absorber concentration for the 4.3-μ (2350-cm^{-1}) CO_2 band.

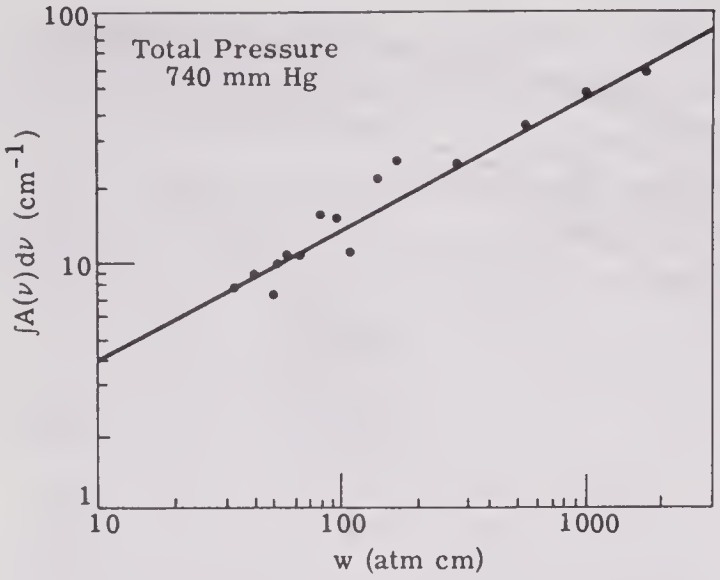


FIG. 6-79. Total absorption *vs.* absorber concentration for the 4.8-μ (2075-cm^{-1}) CO_2 band.

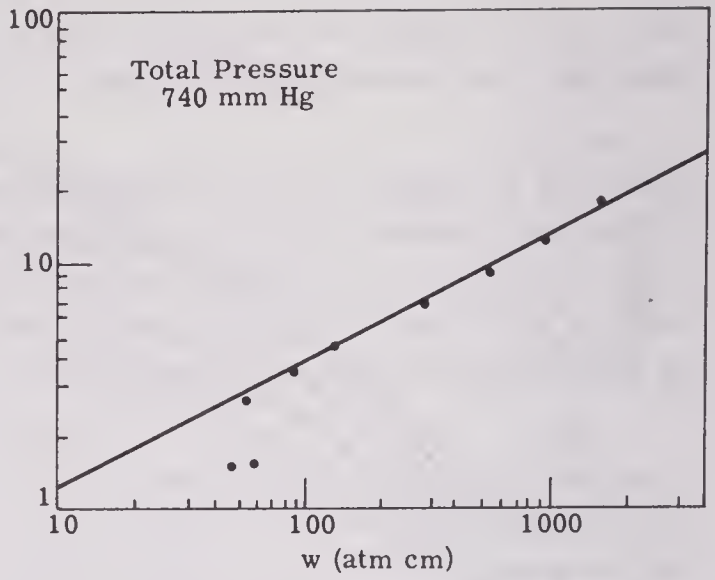


FIG. 6-80. Total absorption *vs.* absorber concentration for the 5.2-μ (1930-cm^{-1}) CO_2 band.

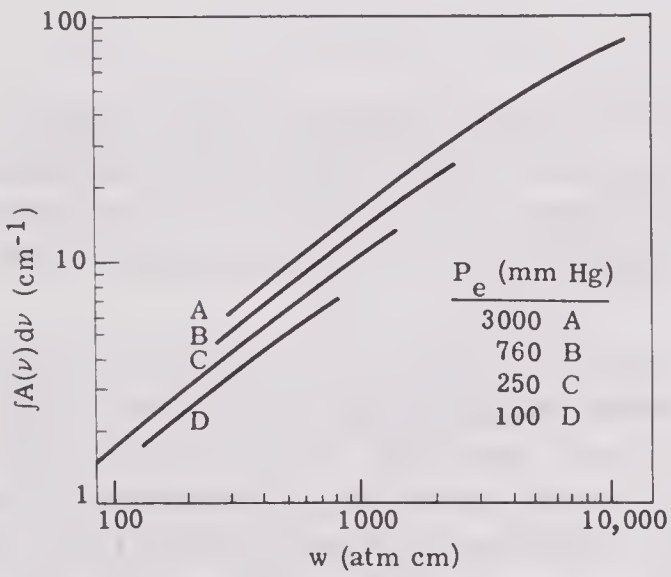


FIG. 6-81. Total absorption *vs.* absorber concentration for the 9.4-μ (1064-cm^{-1}) CO_2 band.

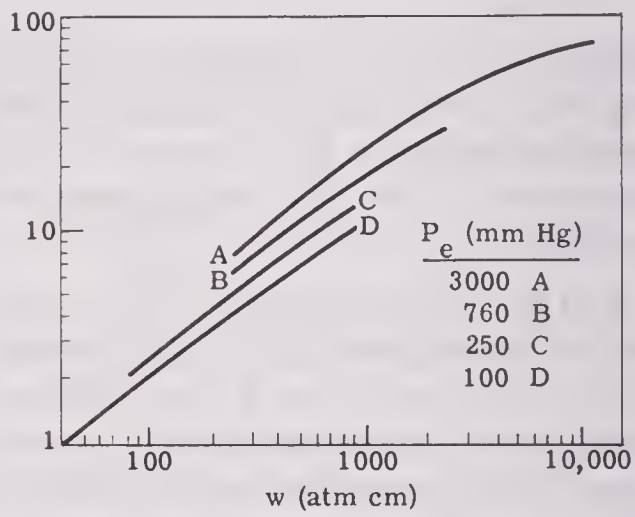


FIG. 6-82. Total absorption *vs.* absorber concentration for the 10.4-μ (961-cm^{-1}) CO_2 band.

6.8.1.3. *The* 9.4-μ (1064-cm^{-1}) *and* 10.4-μ (961-cm^{-1}) *Bands.* Total absorption of the weak CO_2 bands at 9.4 μ (1064 cm^{-1}) and 10.4 μ (961 cm^{-1}) are shown in Fig. 6-81 and 6-82, respectively. Total absorption of these bands is strongly dependent upon temperature; thus, values of total absorption are temperature corrected. The curves shown in Fig. 6-81 and 6-82 are for a temperature of 26°C, and are estimated to be accurate within ±5% for more than 10-cm^{-1} total absorption greater than 10 cm^{-1}, ±10% between 3 and 10-cm^{-1} total absorption, and ±20% below 3 cm^{-1} total absorption.

The temperature dependence of both bands is illustrated in Fig. 6-83. From the curves, it can be seen that total absorption increases with temperature.

6.8.1.4. *The* 1.4-μ (7150-cm^{-1}), 1.6-μ (6250-cm^{-1}), *and* 2.0-μ (5000-cm^{-1}) *Bands.* Total absorption of the very weak CO_2 bands at 1.4 μ (7150 cm^{-1}), 1.6 μ (6250 cm^{-1}) and 2.0 μ (5000 cm^{-1}) are shown in Fig. 6-84 through 6-86. The band at 1.6 μ (6250 cm^{-1}) is actually a group of very weak bands centered at 1.645 μ (6077 cm^{-1}), 1.604 μ (6231 cm^{-1}), 1.574 μ (6351 cm^{-1}), and 1.536 μ (6510 cm^{-1}). The CO_2 band near 2 μ (5000 cm^{-1}) consists primarily of three weak absorption bands centered at 2.057 μ (4861 cm^{-1}), 2.006 μ (4983.5 cm^{-1}), and 1.957 μ (5109 cm^{-1}).

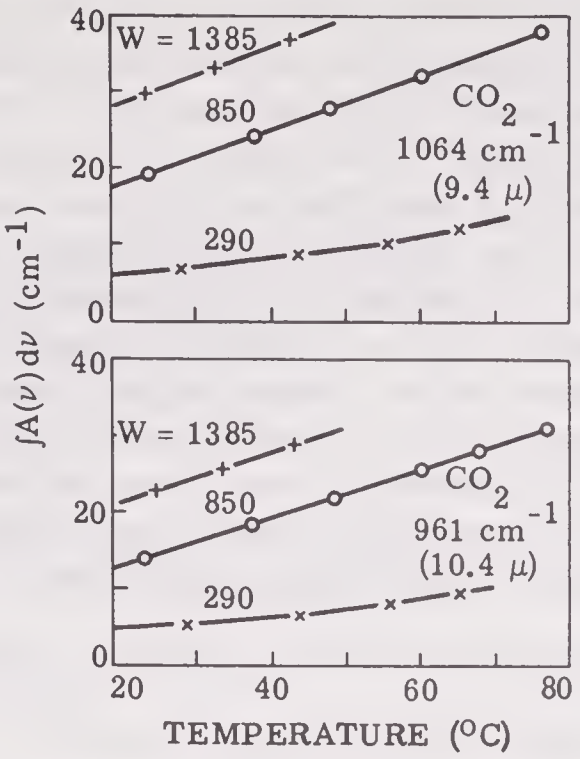


FIG. 6-83. Effects of temperature on total absorption for the 9.4-μ (1064-cm^{-1}) and 10.4-μ (961-cm^{-1}) CO_2 bands.

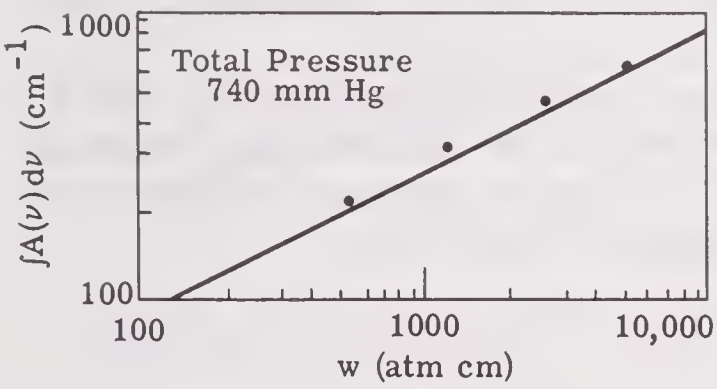


FIG. 6-84. Total absorption *vs.* absorber concentration for the 1.4-μ (7150-cm^{-1}) CO_2 band.

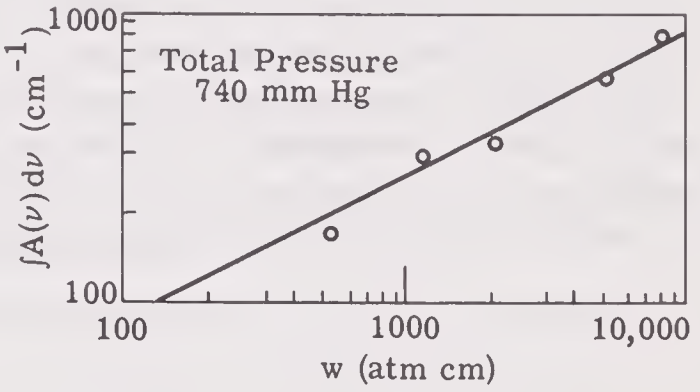


FIG. 6-85. Total absorption *vs.* absorber concentration for the 1.6-μ (6250-cm^{-1}) CO_2 band.

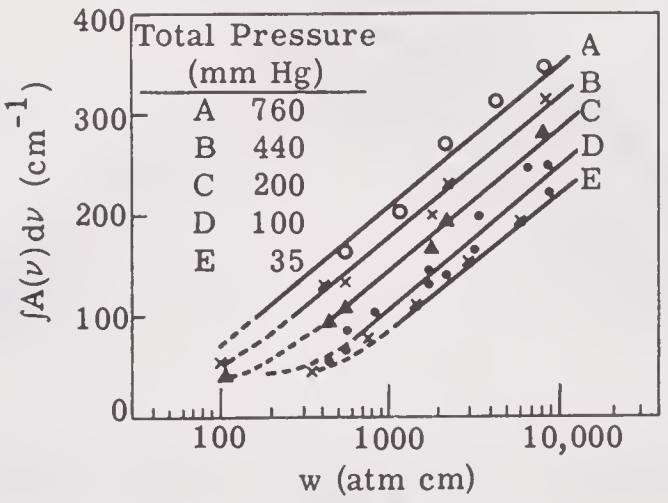


FIG. 6-86. Total absorption *vs.* absorber concentration for the 2-μ (5000-cm^{-1}) CO_2 band.

6.8.1.5. *The* 11.4-μ (875-cm^{-1}) *to* 20-μ (495-cm^{-1}) *Region.* There are several strong and medium CO_2 absorption bands in the spectral region between 11.4 μ (875 cm^{-1}) and 20 μ (495 cm^{-1}), the strongest of which is centered at 14.9 μ (667 cm^{-1}). Because of the wide spectral region covered in this section, the spectral region is divided into smaller subregions, and absorber concentration is plotted against mean fraction absorption rather than against total absorption, which usually refers to an entire band. Total absorption $\int A(\nu)d\nu$ of an entire band is independent of spectral slit width, provided that there is no absorption beyond the limits of integration. It follows that the total absorption of a subregion will be independent of spectral slit width if the subregions are divided at frequencies where the absorption is zero. Practically, however, no frequencies exist where the absorption is zero, and optimum frequencies, *i.e.,* frequencies at which absorption is very slight, are chosen to divide the subregions.

Mean fractional absorption is related to total absorption by the following equation:

$$\overline{A}(\nu_1 - \nu_2) = \frac{1}{(\nu_2 - \nu_1)} \int_{\nu_1}^{\nu_2} A(\nu)\, d\nu \tag{6-53}$$

which represents the mean fractional absorption in the spectral region $\nu_1 - \nu_2$.

Figures 6-87 through 6-91 show mean fractional absorption for each of the five subregions of the 11.4-μ (875-cm^{-1}) to 20.2 μ (495-cm^{-1}) CO_2 region. The curves are based upon a temperature of 26°C, and are estimated to be accurate within ±5% for a mean fractional absorption greater than 0.10 cm^{-1}, and increase to approximately ±20% for smaller values of total absorption.

The effect of temperature on mean fractional absorption is shown in Fig. 6-92 for four of the five subregions. No data are available for the fifth subregion, which extends from 18.3 μ (545 cm^{-1}) to 20 μ (495 cm^{-1}). From the curves in Fig. 6-92, it can be seen that mean fractional absorption increases with temperature.

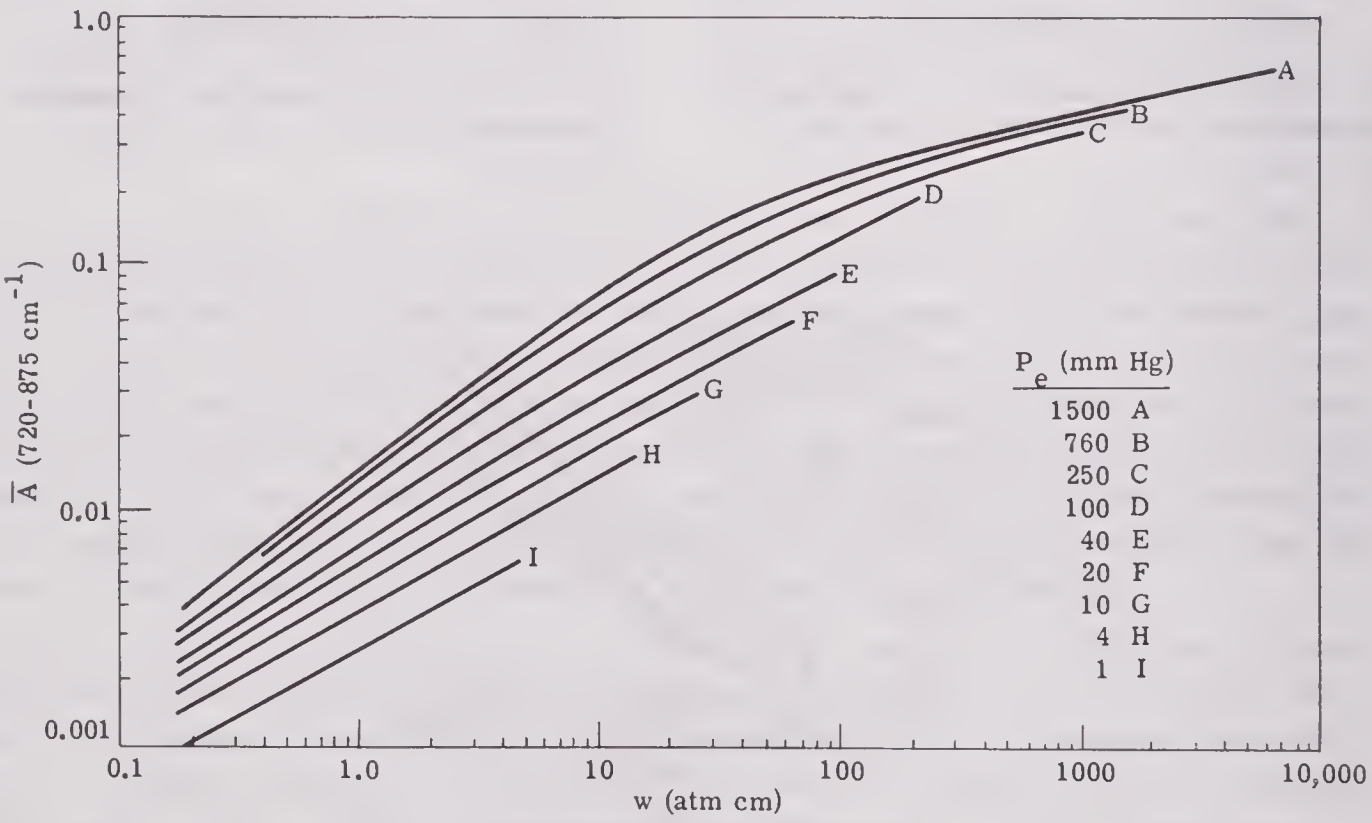


FIG. 6-87. Mean fractional absorption *vs.* absorber concentration for 11.4-μ (875-cm^{-1}) to 13.9-μ (720-cm^{-1}) CO_2 band.

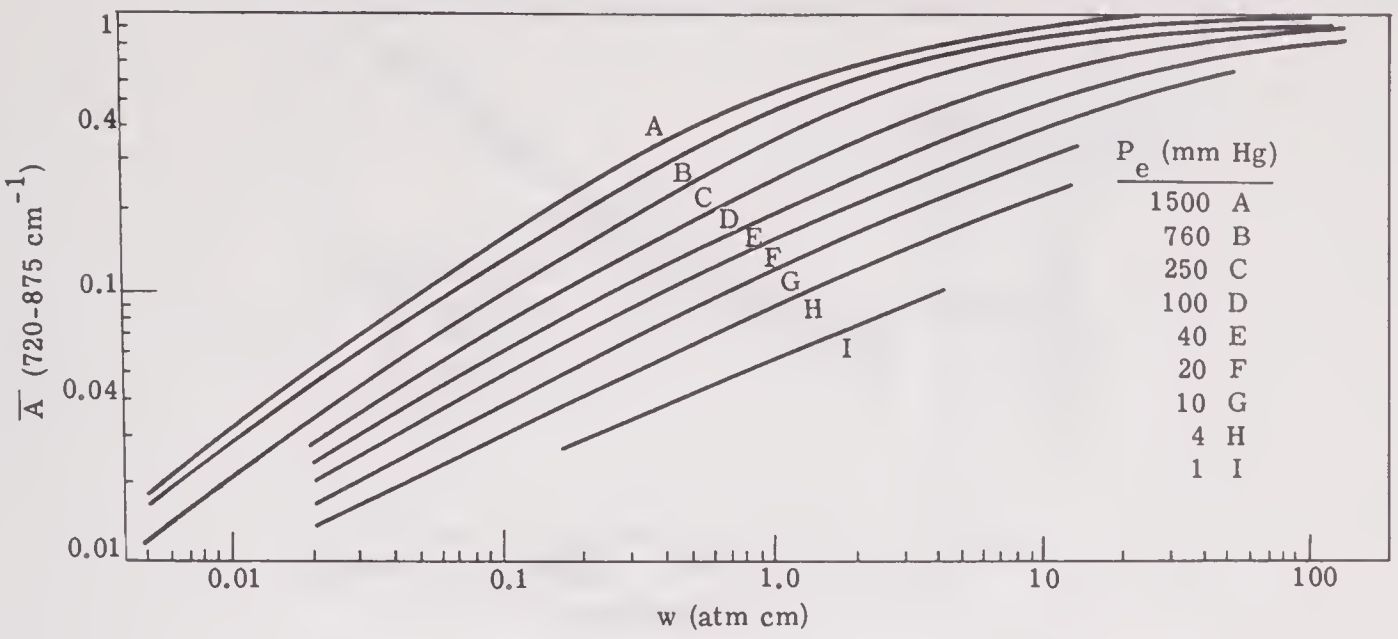


FIG. 6-88. Mean fractional absorption *vs.* absorber concentration for the 13.9-μ (720-cm^{-1}) to 14.9-μ (667-cm^{-1}) CO_2 band.

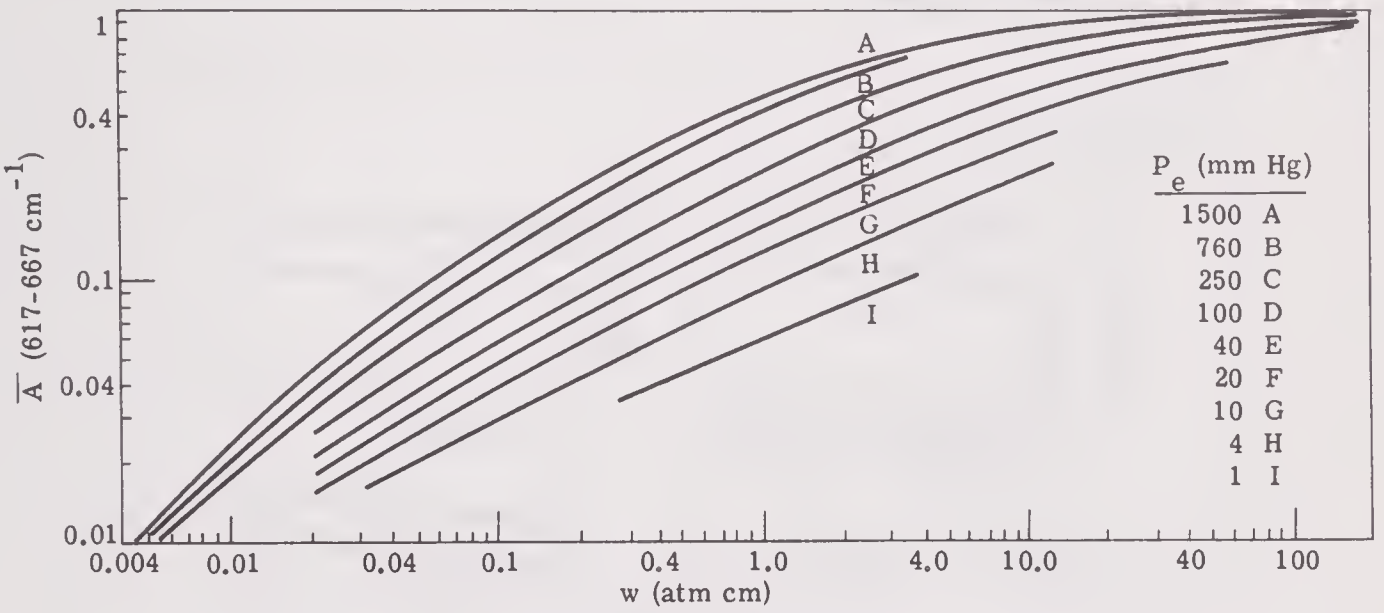


FIG. 6-89. Mean fractional absorption *vs.* absorber concentration for the 14.9-μ (667-cm^{-1}) to 16.2-μ (617-cm^{-1}) CO_2 band.

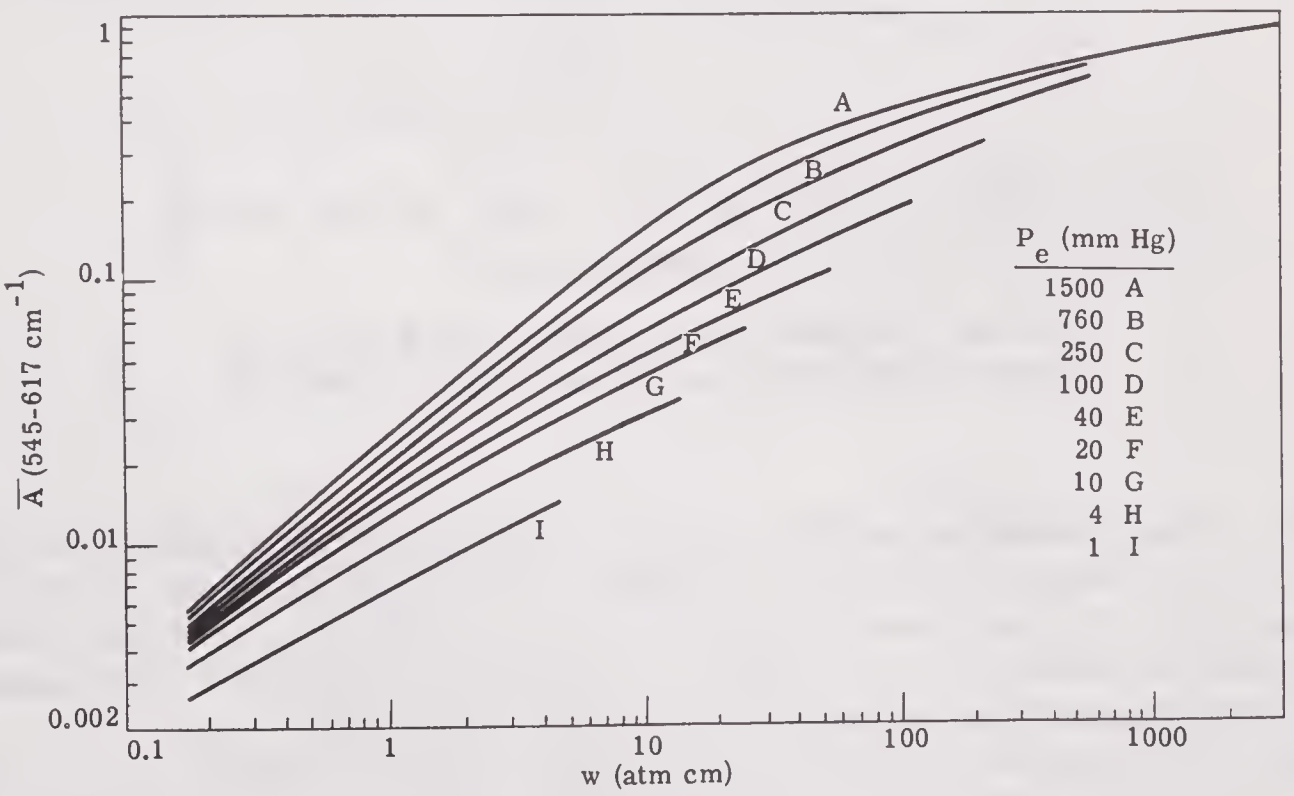


FIG. 6-90. Mean fractional absorption *vs.* absorber concentration for the 16.2-μ (617-cm^{-1}) to 18.3-μ (545-cm^{-1}) CO_2 band.

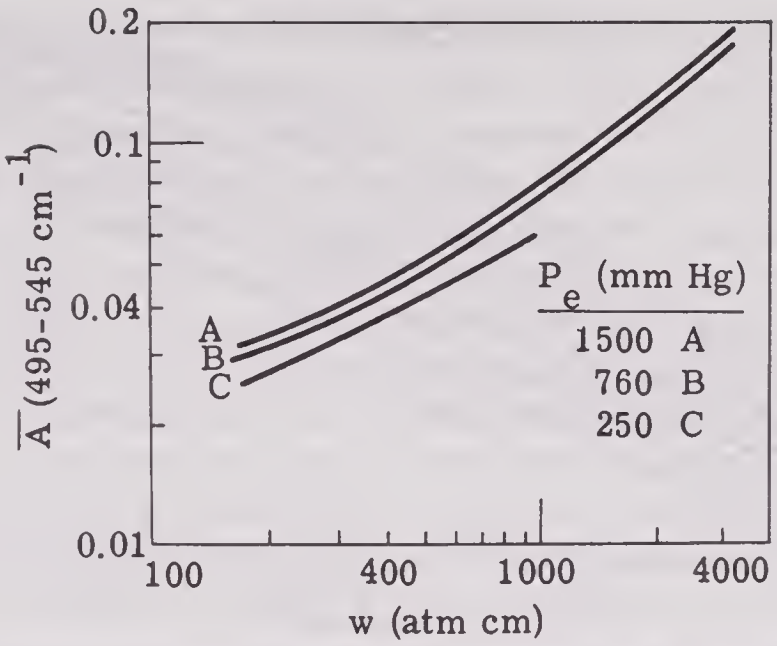


FIG. 6-91. Mean fractional absorption *vs.* absorber concentration for the 18.3-μ (545-cm^{-1}) to 20-μ (495-cm^{-1}) CO_2 band.

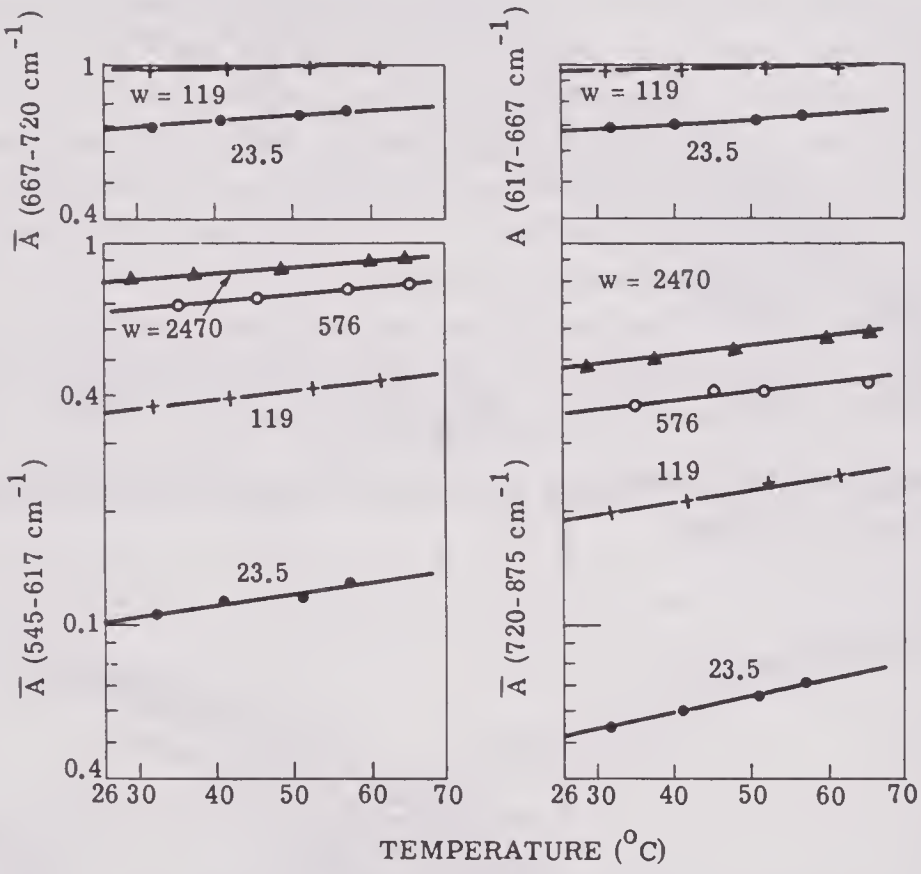


FIG. 6.92. Effects of temperature on mean fractional absorption of CO_2 from 11.4 μ (875 cm^{-1}) to 18.3 μ (545 cm^{-1}).

6.8.2. Total Absorption by H_2O [106, 107]. Total absorption of the 6.27-μ (1595-cm^{-1}), 2.70-μ (3700-cm^{-1}), and 1.87-μ (5332-cm^{-1}) H_2O bands are shown in Fig. 6-93 through 6-95. The absorption in the 2.70-μ region is caused primarily by two absorption bands with centers at 2.73 μ (3657 cm^{-1}) and 2.66 μ (3756 cm^{-1}). The curves in Fig. 6-93 through 6-95 are estimated to be accurate within ±6%.

Minor H_2O absorption bands at 3.2 μ (3150 cm^{-1}), 1.38 μ (7250 cm^{-1}), 1.1 μ (8807 cm^{-1}), and 0.94 μ (10,613 cm^{-1}) are shown in Fig. 6-96 through 6-99. These curves are estimated to be accurate within about ±10%.

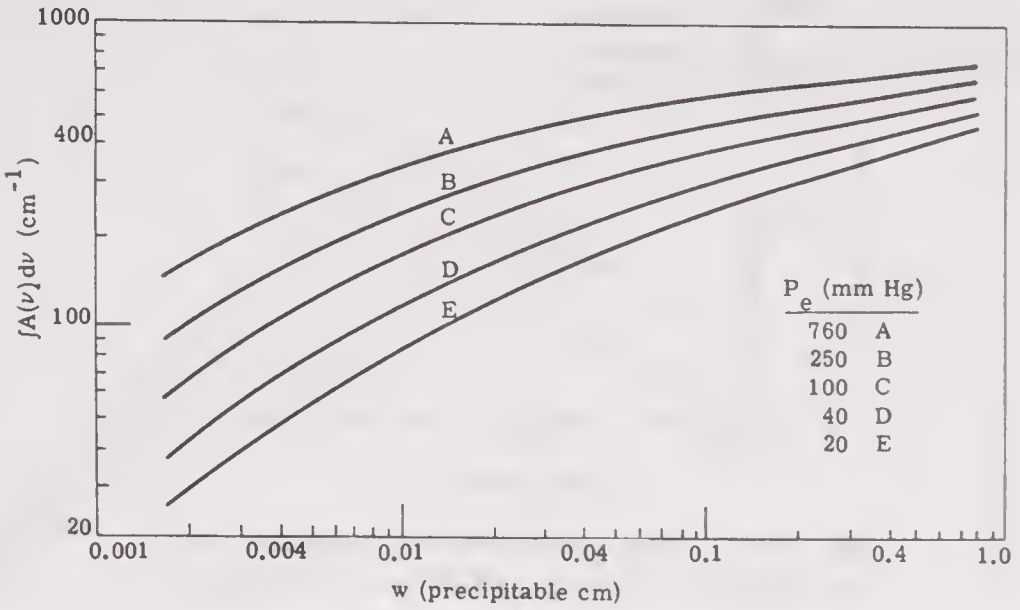


FIG. 6-93. Total absorption *vs.* absorber concentration for the 6.27-μ (1595-cm^{-1}) H_2O band.

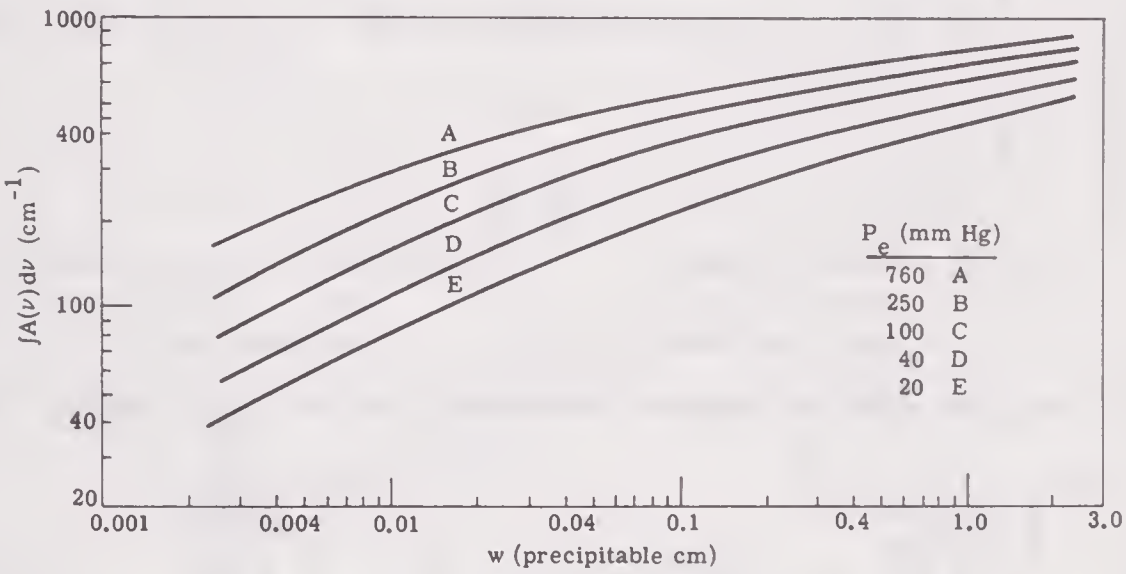


FIG. 6-94. Total absorption *vs.* absorber concentration for the 2.70-μ (3700-cm^{-1}) H_2O band.

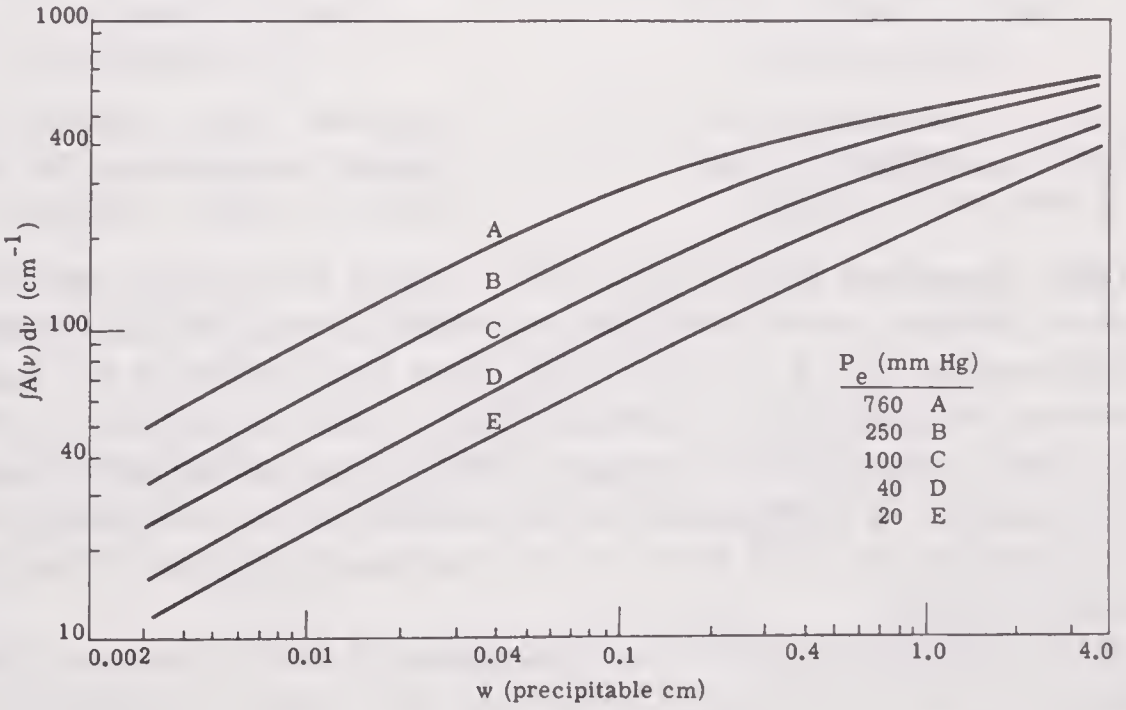


FIG. 6-95. Total absorption *vs.* absorber concentration for the 1.87-μ (5332-cm^{-1}) H_2O band.

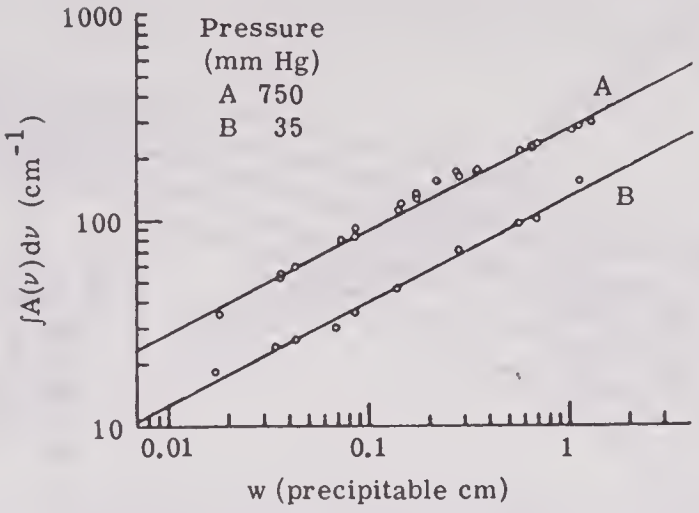


FIG. 6-96. Total absorption *vs.* absorber concentration for the 3.2-μ (3150-cm^{-1}) H_2O band.

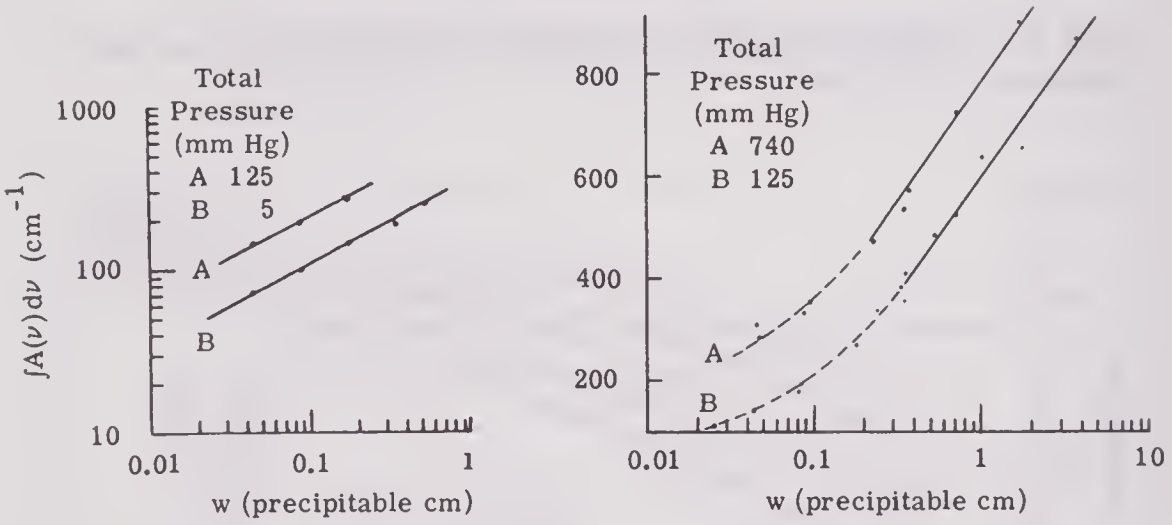


FIG. 6-97. Total absorption *vs.* absorber concentration for the 1.38-μ (7250-cm^{-1}) H_2O band.

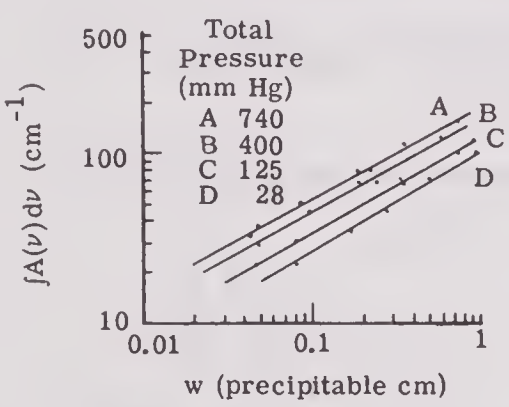


FIG. 6-98. Total absorption *vs.* absorber concentration for the 1.1-μ (8807-cm^{-1}) H_2O band.

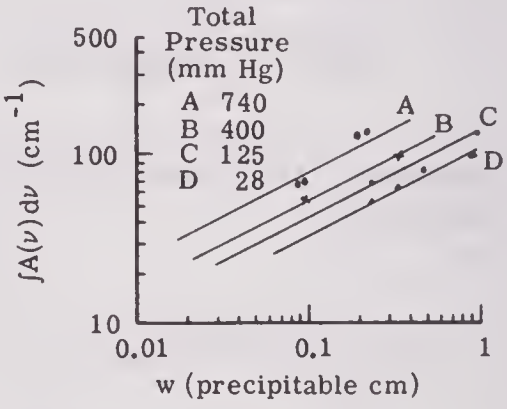


FIG. 6-99. Total absorption *vs.* absorber concentration for the 0.94-μ (10,163-cm^{-1}) H_2O band.

6.8.3. Total Absorption by N_2O [106, 107]. Figure 6-100 shows total absorption as a function of absorber concentration and equivalent pressure for the ν_3 fundamental band of N_2O centered at 4.5 μ (2224 cm^{-1}). This band occurs in the atmospheric window between the strong 4.3-μ (2350-cm^{-1}) CO_2 band and the 6.2-μ (1595-cm^{-1}) H_2O band. Thus, it gives rise to the major portion of the atmospheric absorption in the 4.3-6.2-μ window. It is estimated that the accuracy of the total absorption values given in Fig. 6-100 is within ±5% below 10 cm^{-1} total absorption and within ±5% above 10 cm^{-1} total absorption.

Figures 6-101 through 6-105 show total absorption of N_2O in the 3.9-μ (2563-cm^{-1}), 4.05-μ (2461-cm^{-1}), 7.7-μ (1285-cm^{-1}), 8.6-μ (1167 cm^{-1}), and 17.1-μ (589 cm^{-1}) bands. These curves show total absorption within an estimated accuracy of ±5% above a total absorption of 30 cm; below 30 cm^{-1}, the accuracy is somewhat less.

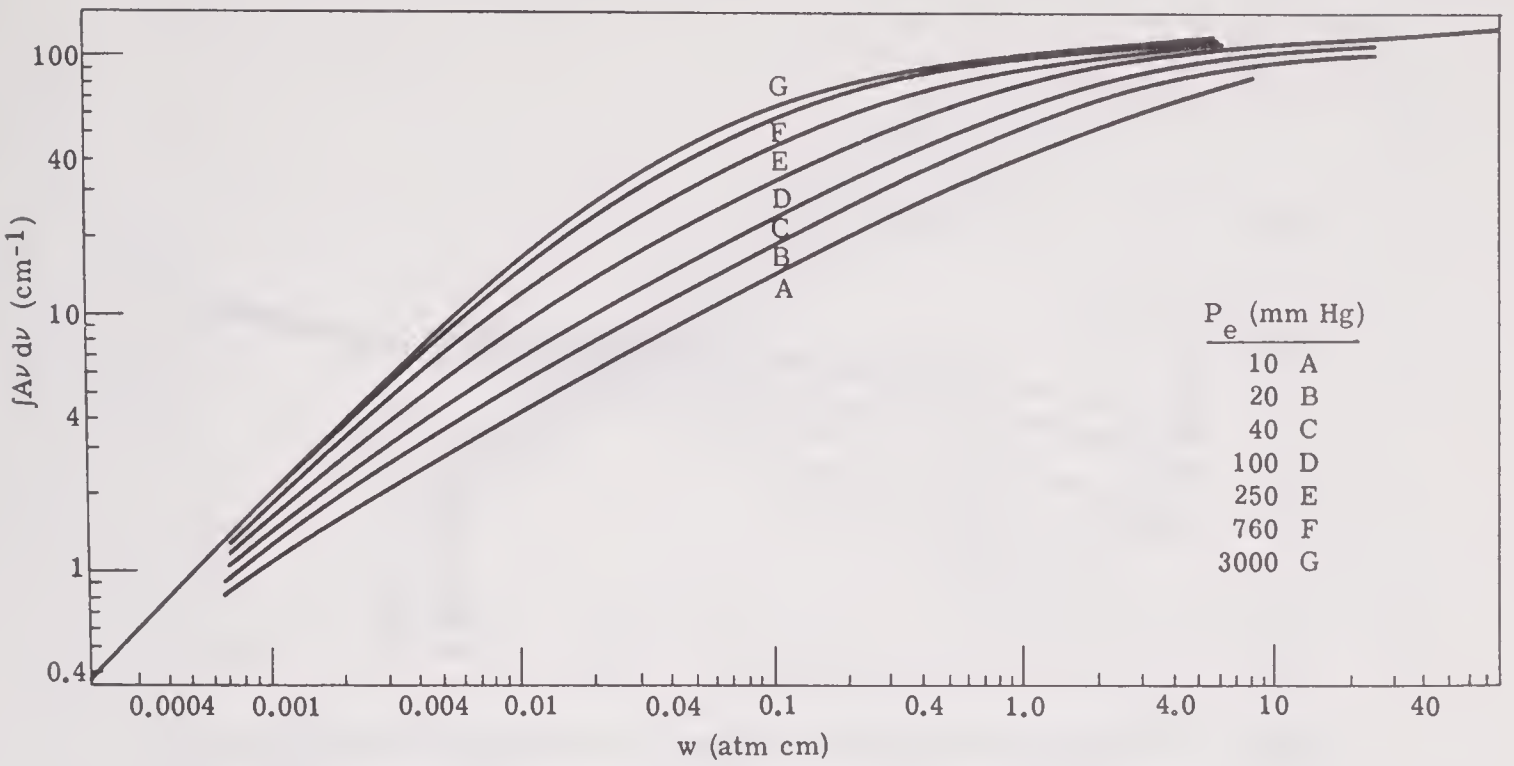


FIG. 6-100. Total absorption *vs.* absorber concentration for the 4.5-μ (2224-cm^{-1}) N_2O band.

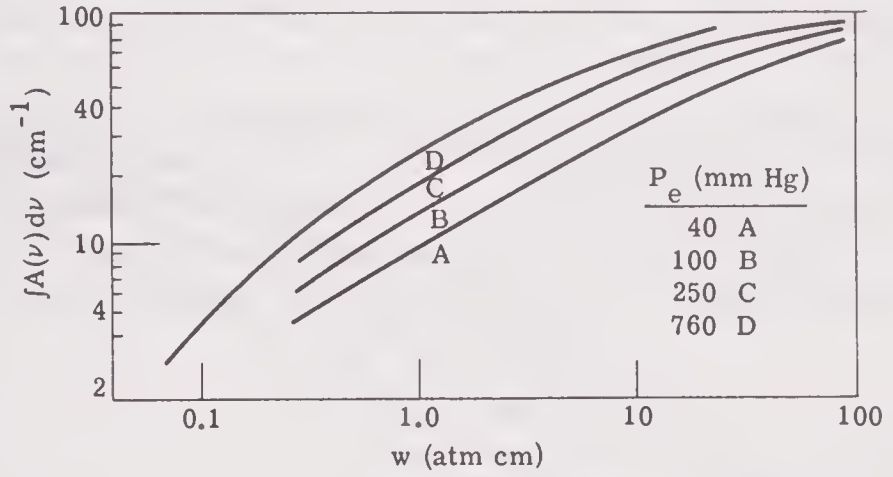


FIG. 6-101. Total absorption *vs.* absorber concentration for the 3.9-μ (2563-cm^{-1}) N_2O band.

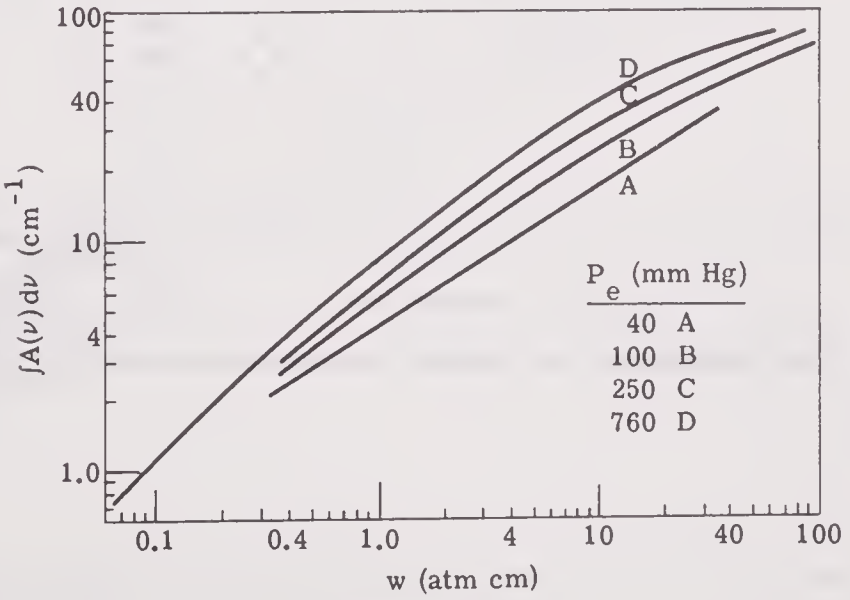


FIG. 6-102. Total absorption *vs.* absorber concentration for the 4.05-μ (2461-cm^{-1}) N_2O band.

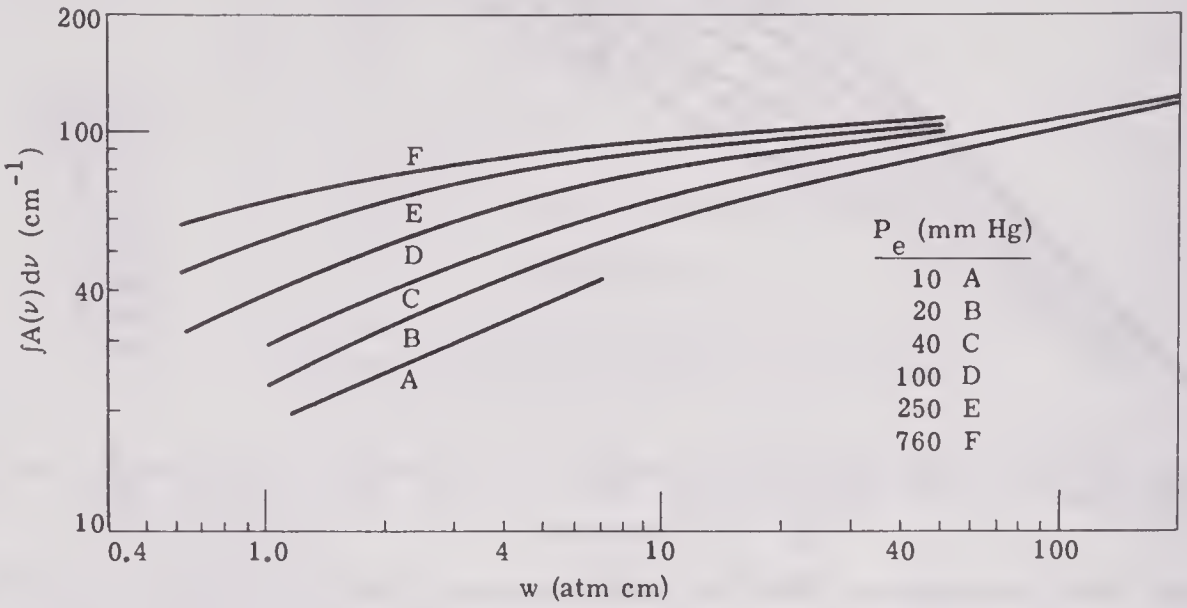


FIG. 6-103. Total absorption *vs.* absorber concentration for the 7.7-μ (1285-cm^{-1}) N_2O band.

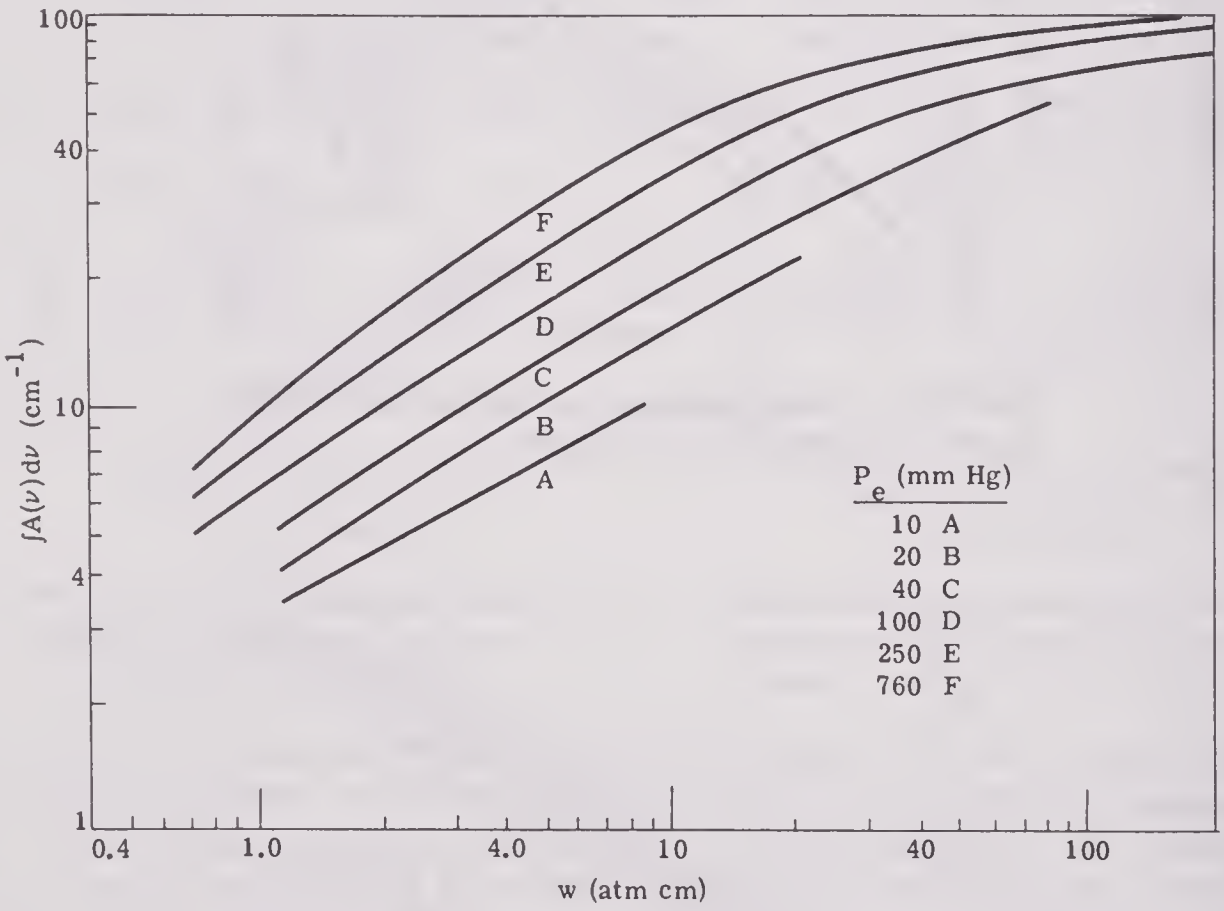


FIG. 6-104. Total absorption *vs.* absorber concentration for 8.6-μ (1167-cm^{-1}) N_2O band.

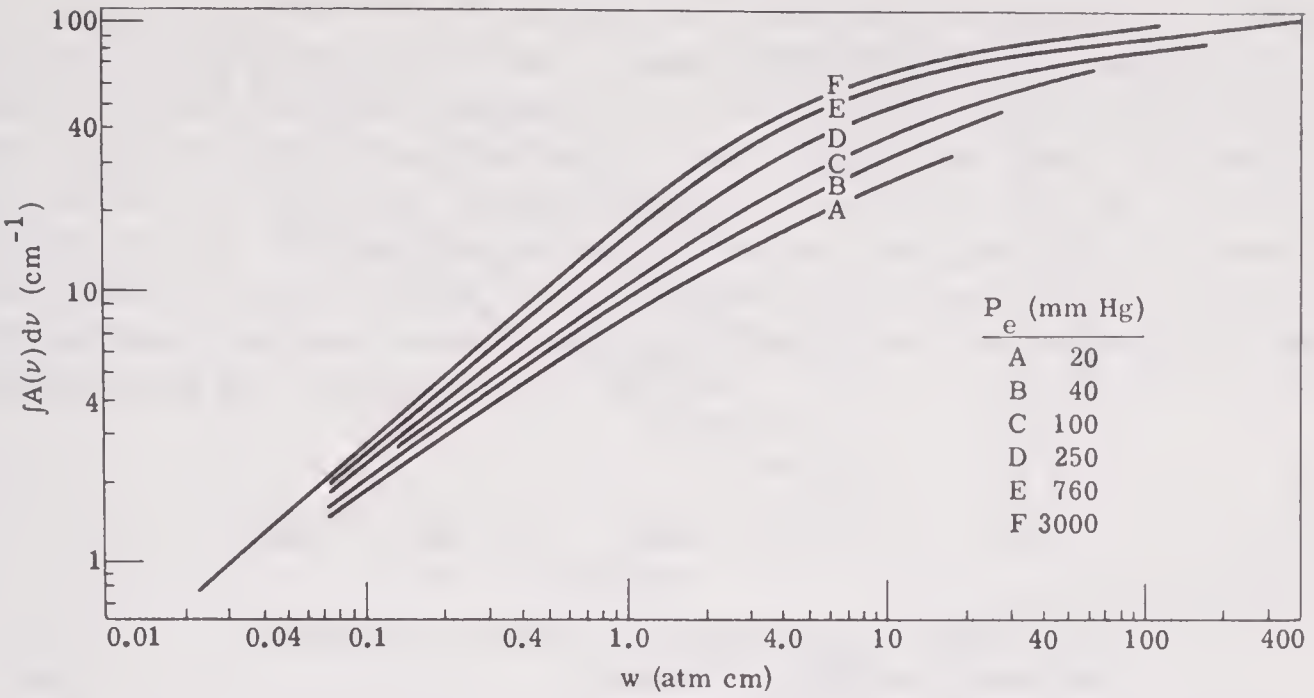


FIG. 6-105. Total absorption *vs.* absorber concentration for the 17.1-μ (589-cm^{-1}) N_2O band.

6.8.4. Total Absorption by CO. Total absorption of the 4.6-μ (2143-cm^{-1}) CO band is shown in Fig. 6-106. The data are estimated to be accurate within ±5% above 10-cm^{-1} total absorption and within ±10% below 10-cm^{-1} total absorption.

Figure 6-107 shows total absorption for the 2.3-μ (4260-cm^{-1}) CO band. The curves in Fig. 6-107 do not represent nearly as wide a range of absorber concentration and equivalent pressure as those of Fig. 6-106 because the largest values of equivalent pressure were not sufficiently large, nor the smallest values of absorber concentration sufficiently small, to produce complete absorption over the bands. By comparing Fig. 6-106 and 6-107, however, it can be seen that, at any given equivalent pressure, the total

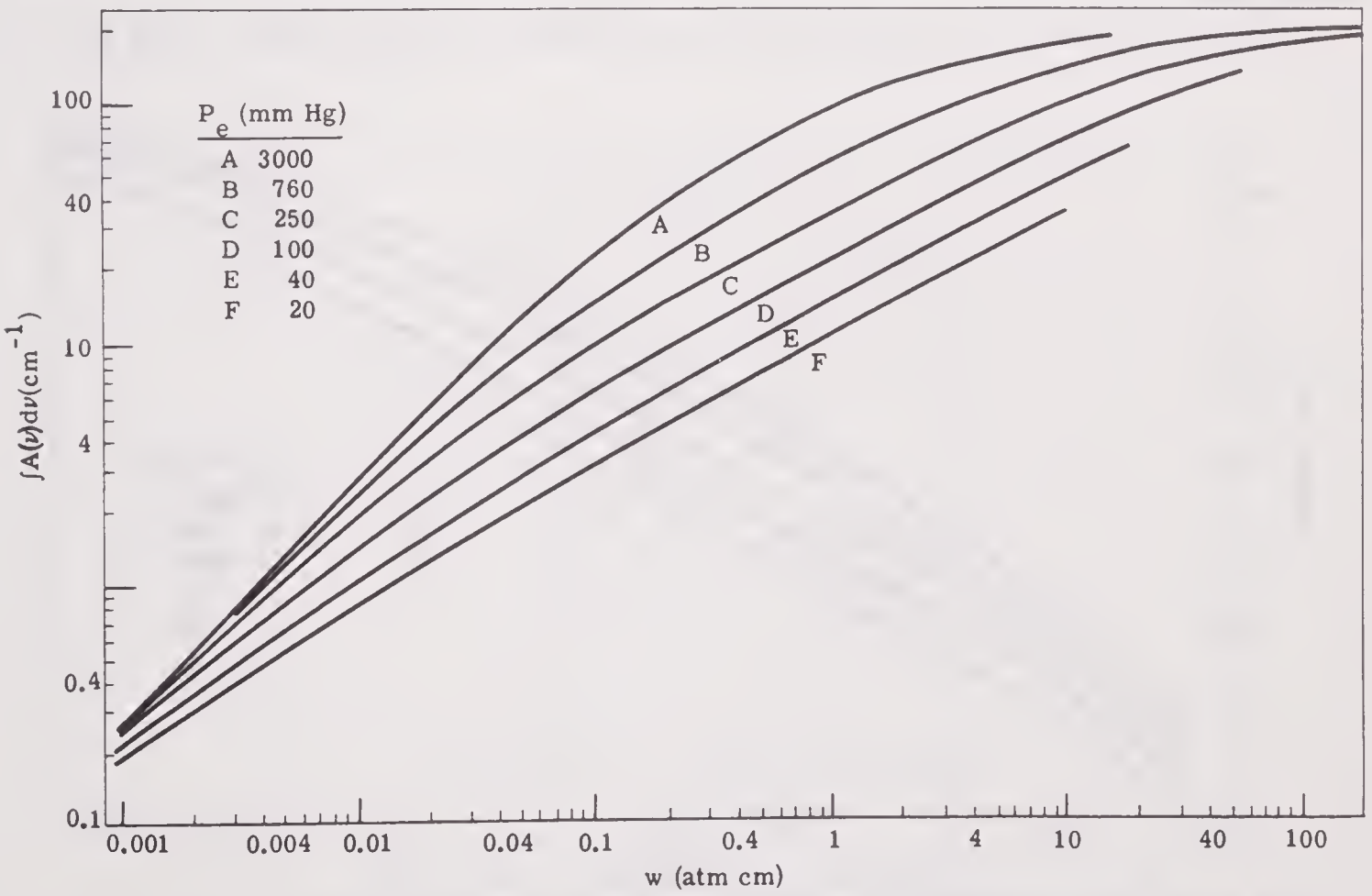


FIG. 6-106. Total absorption *vs.* absorption concentration for the 4.6-μ (2143-cm^{-1}) CO band.

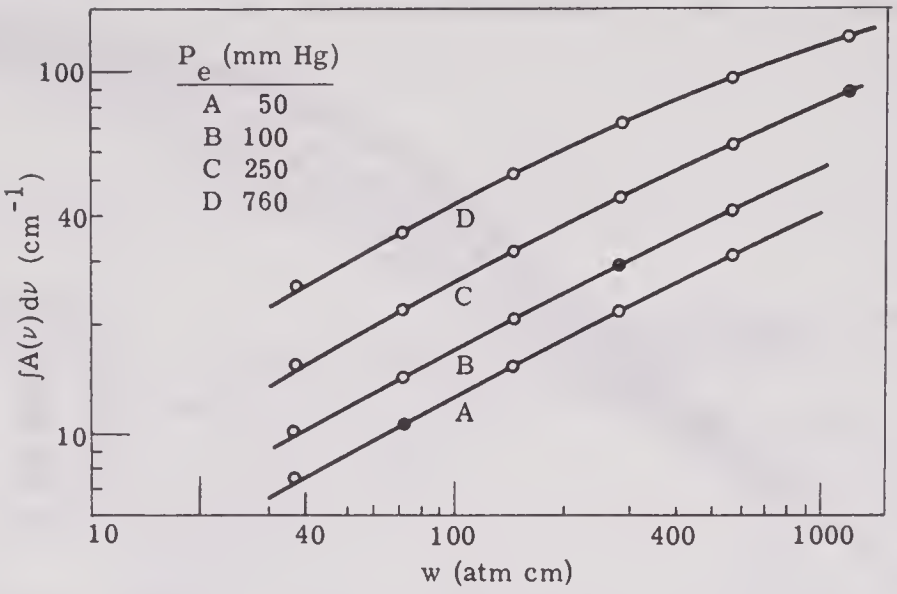


FIG. 6-107. Total absorption *vs.* absorber concentration for the 2.3-μ (4260-cm^{-1}) CO band.

absorption of the 4.6-μ (2143-cm^{-1}) band is very nearly the same as that of the 2.3-μ (4260-cm^{-1}) band at the same equivalent pressure but with the value of the absorber concentration 150 times as great. Thus, since the line shape, line spacing, and relative line strength within the bands are similar, the total absorption of the weaker overtone band (Fig. 6-107) can probably be determined from the curves of the fundamental band (Fig. 6-106) by use of the same value of equivalent pressure and a value of absorber concentration that is 1/150 that of the absorber whose total absorption is being measured.

6.8.5. Total Absorption by CH_4. Figure 6-108 shows the total absorption of the 3.31-μ (3020-cm^{-1}) CH_4 band. The values of total absorption are estimated to be accurate within ±5% above 10-cm^{-1} total absorption and within ±10% below 10-cm^{-1} total absorption.

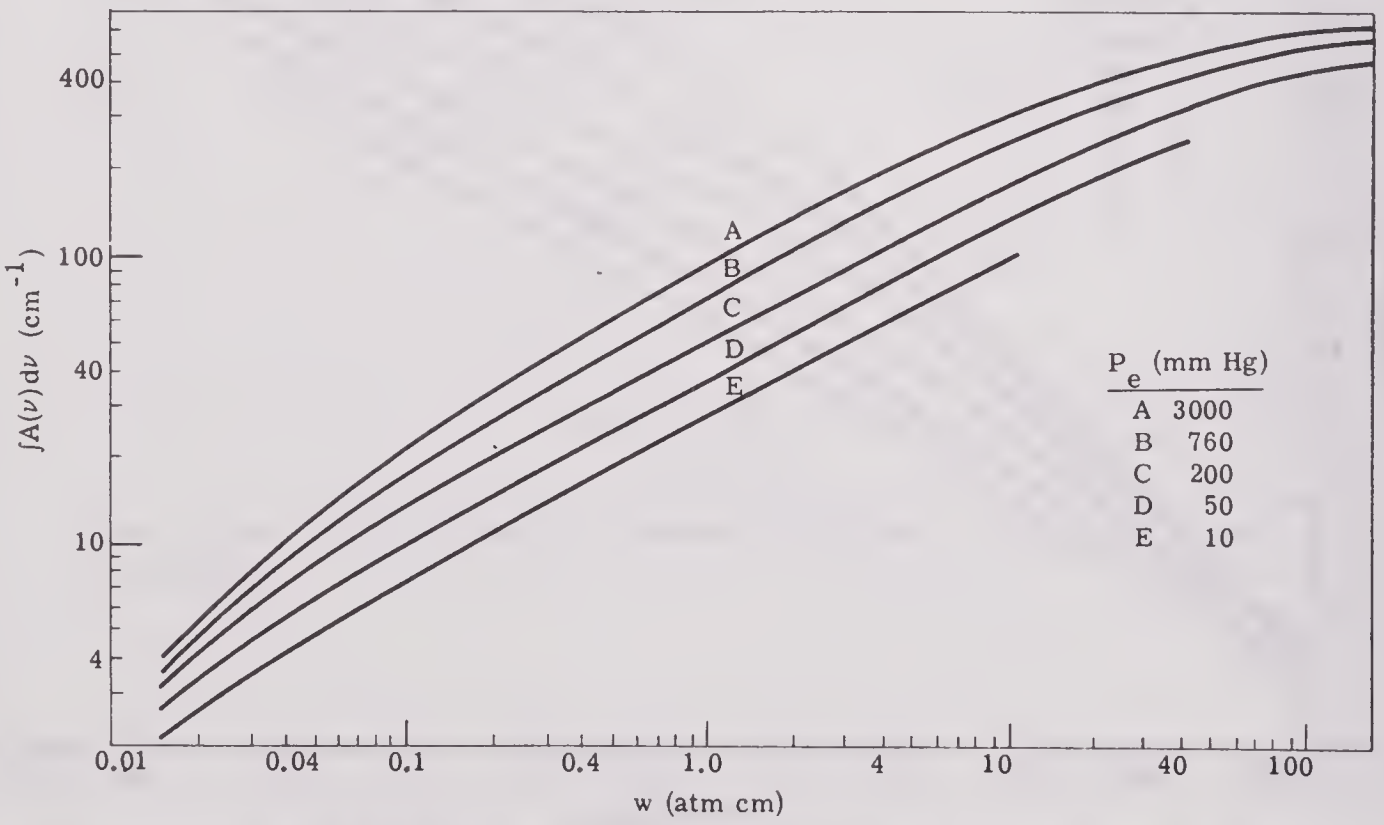


FIG. 6-108. Total absorption *vs.* absorber concentration for the 3.3-μ (3020-cm^{-1}) CH_4 band.

Figures 6-109 and 6-110 show the total absorption of the 7.6-μ (1306-cm^{-1}) CH_4 band and the 6.5-μ (1550-cm^{-1}) CH_4 band, respectively. Actually, absorption by the 7.6-μ (1306-cm^{-1}) band overlaps the absorption by the 6.5-μ (1550-cm^{-1}) band for high absorber concentrations. The estimated accuracy of the total absorption given for the 7.6-μ (1306-cm^{-1}) band (Fig. 6-109) is the same as that for the 3.31-μ (3020-cm^{-1}) band (Fig. 6-108). Because of overlapping by water vapor, however, the curves for the 6.5-μ (1550-cm^{-1}) band (Fig. 6-110) are estimated to be accurate to no better than ±10% for total absorption values greater than about 30 cm^{-1} and ±20% for total absorption values less than about 30 cm^{-1}.

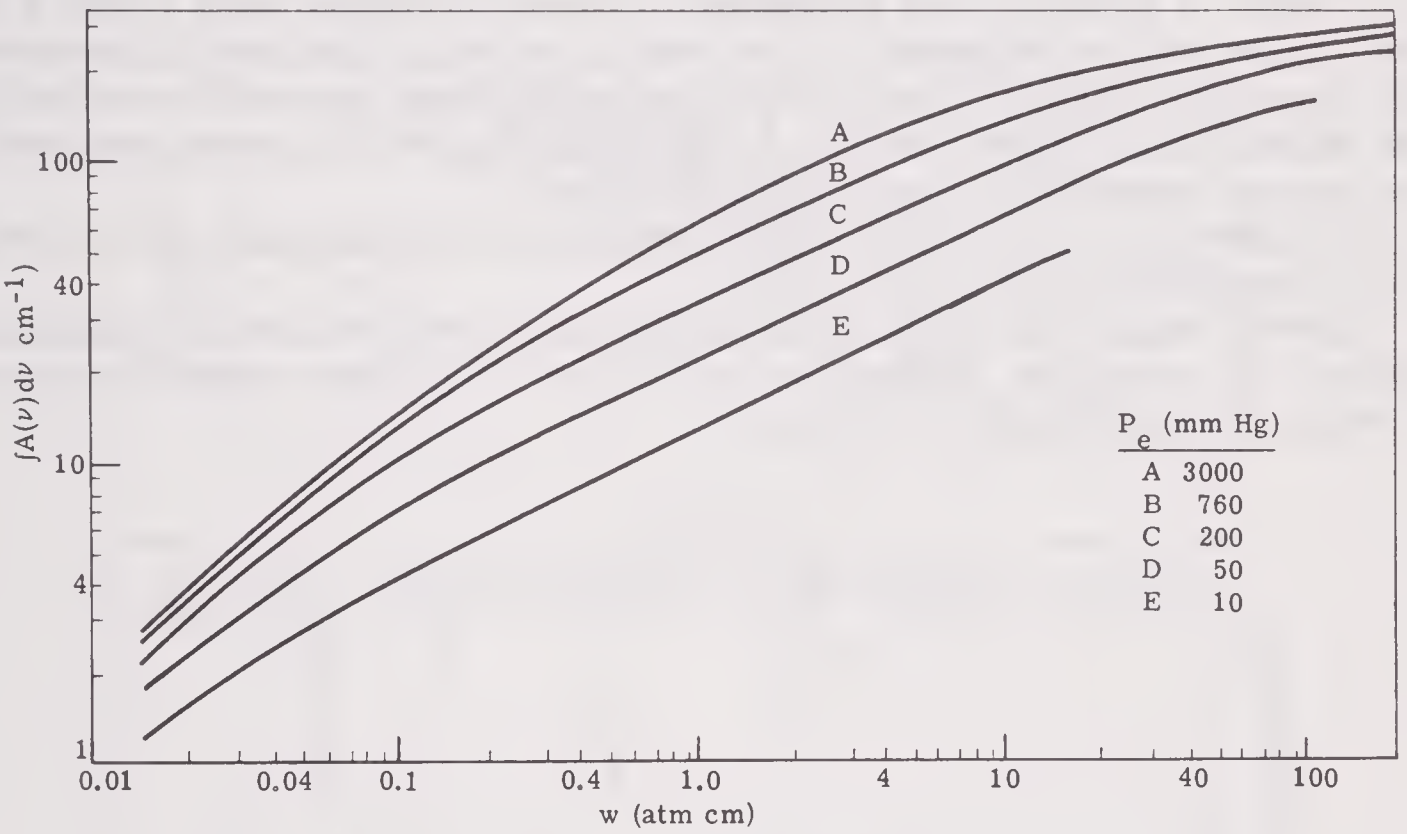


FIG. 6-109. Total absorption *vs.* absorber concentration for the 7.6-μ (1306-cm^{-1}) CH_4 band.

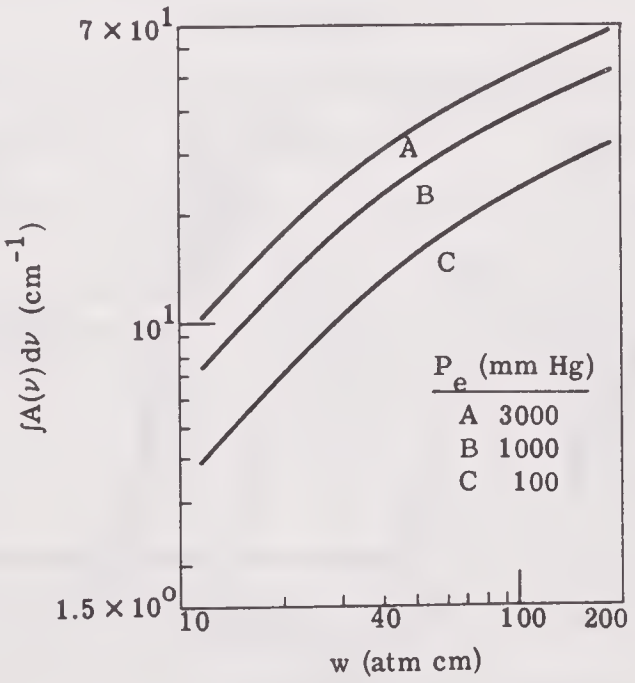


FIG. 6-110. Total absorption *vs.* absorber concentration for the 6.5-μ (1550-cm^{-1}) CH_4 band.

6.9. Infrared Transmission Through the Atmosphere

6.9.1. Horizontal Paths [108]. Sea-level transmission measurements over paths of 0.3, 5.5, and 16.25 km are shown in Fig. 6-111 through 6-119. The 0.3-km path measurements cover the spectral range from approximately 0.5 to 26 μ with a resolution ranging from 1 to 2 wave numbers in the regions beyond 2 μ. The 5.5- and 16.5-km paths cover the spectral range from approximately 0.5 to 15 μ with an average resolving power, $\lambda/\Delta\lambda$, of about 300.

Transmission measurements at 10,000 ft over a path of 27.25 km are shown in Fig. 6-120 and 6-121. The spectral range from approximately 0.5 to 15 μ is covered with an average resolving power, $\lambda/\Delta\lambda$, of about 300.

"Selective-window transmission" from 0.94 to 15 μ is shown in Fig. 6-122. A window is a region of relatively high transmission between absorption bands (Table 6-7). For a given window, transmission is the ratio of the energy that actually penetrates the atmosphere between the limits of the window to the energy that would be received in the absence of any selective absorbers. Thus, it accounts for only selective absorption, and does not take scattering losses into consideration.

The dashed lines in Fig. 6-122 represent the slope of the approximate transmission curves described in [109].

Windows I and IX (Table 6-7) are not plotted because window I is only slightly dependent on water-vapor concentration and window IX becomes essentially opaque through only 5 or 6 mm of water vapor.

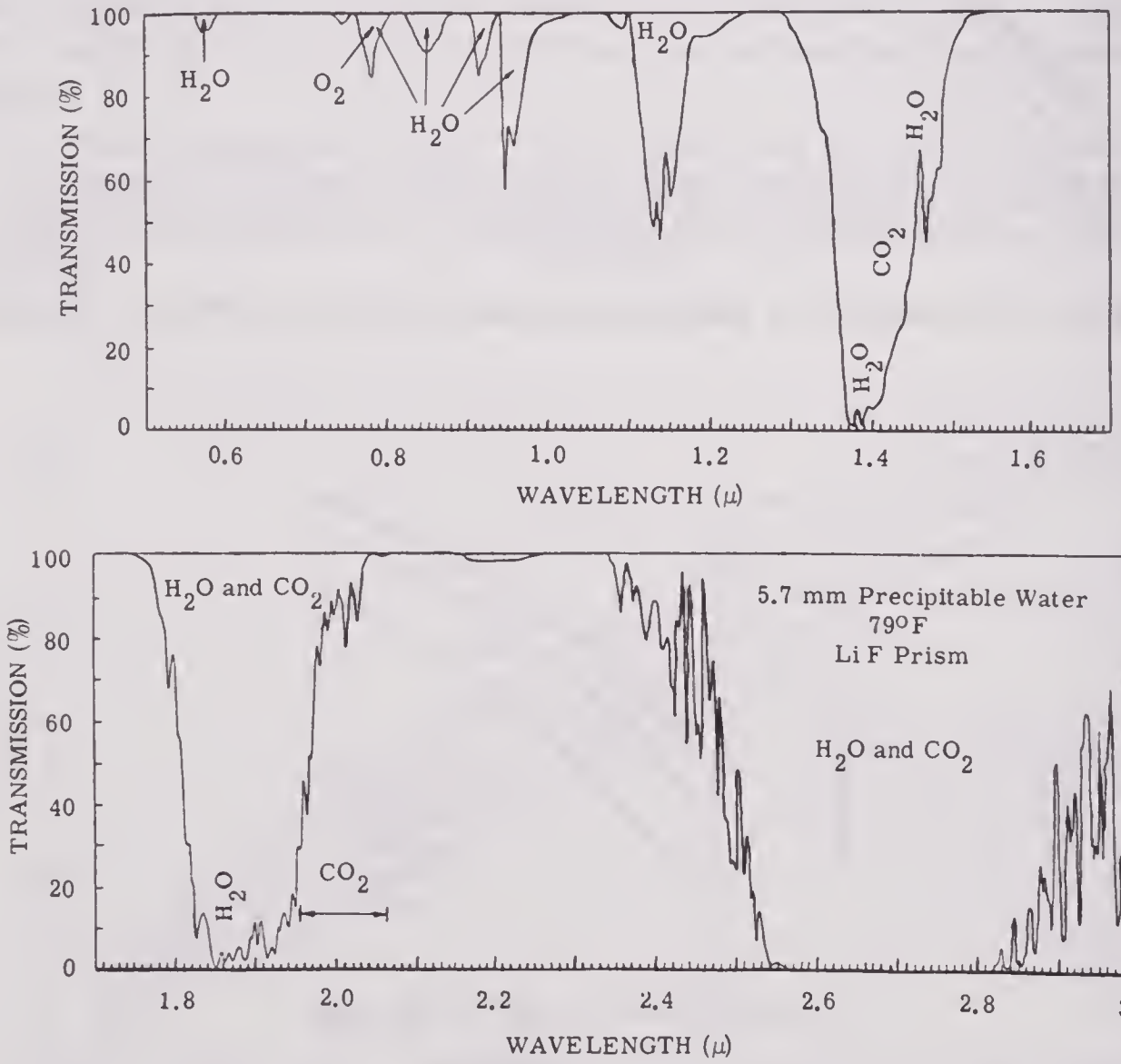


FIG. 6-111. Atmospheric transmission at sea level over a 0.3-km path, 0.5 to 2.8 μ [108].

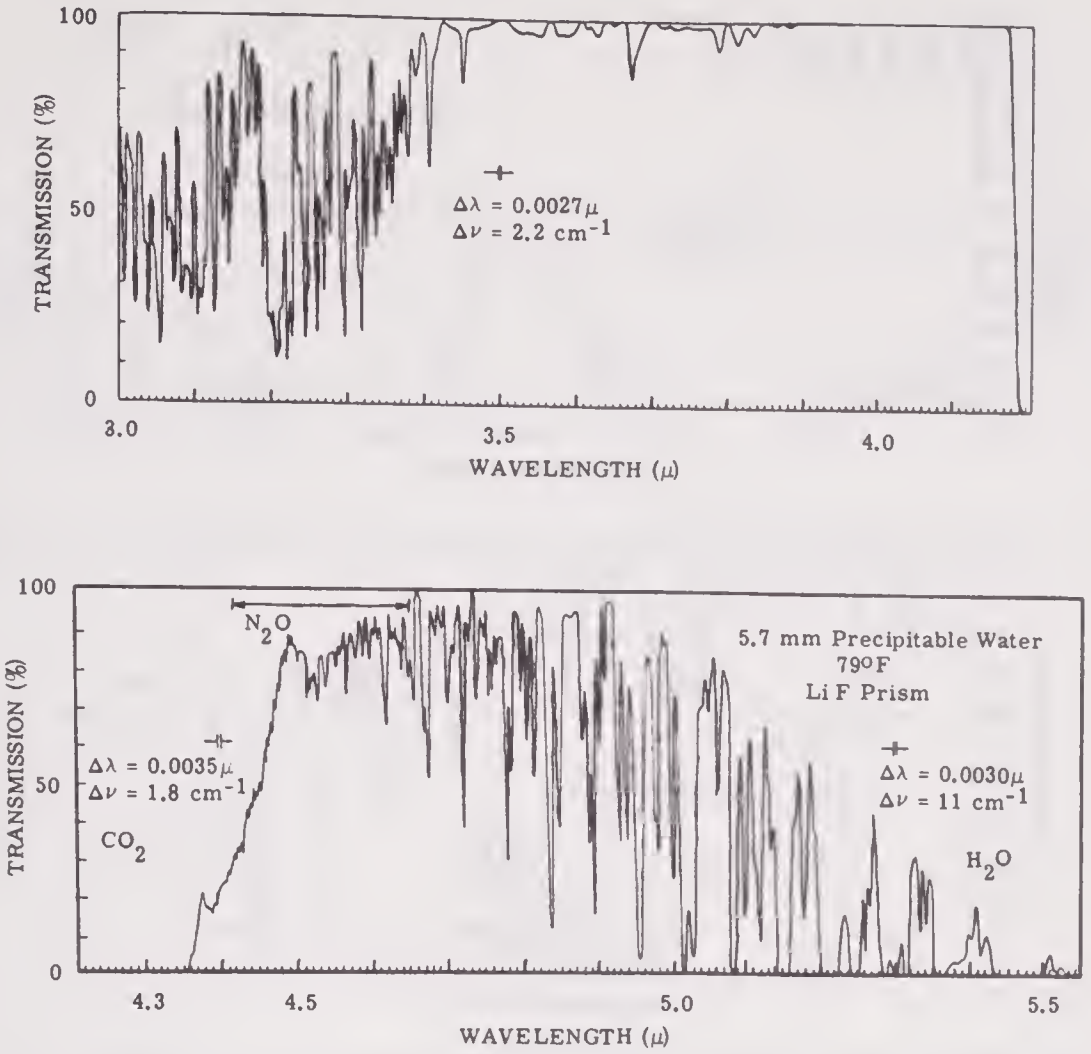


FIG. 6-112. Atmospheric transmission at sea level over a 0.3-km path, 3.0 to 518.5 μ [108].

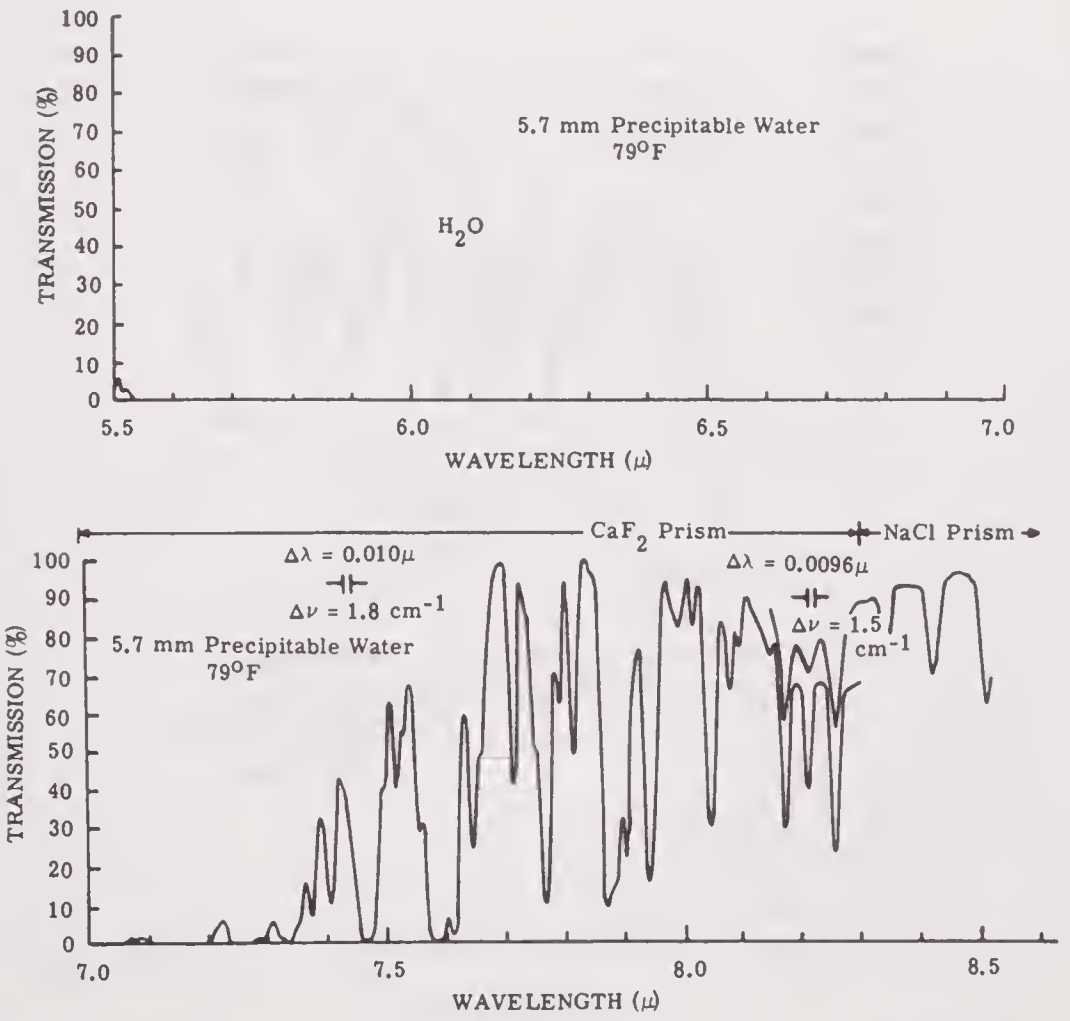


FIG. 6-113. Atmospheric transmission at sea level over a 0.3-km path, 5.5 to 8.5 μ [108].

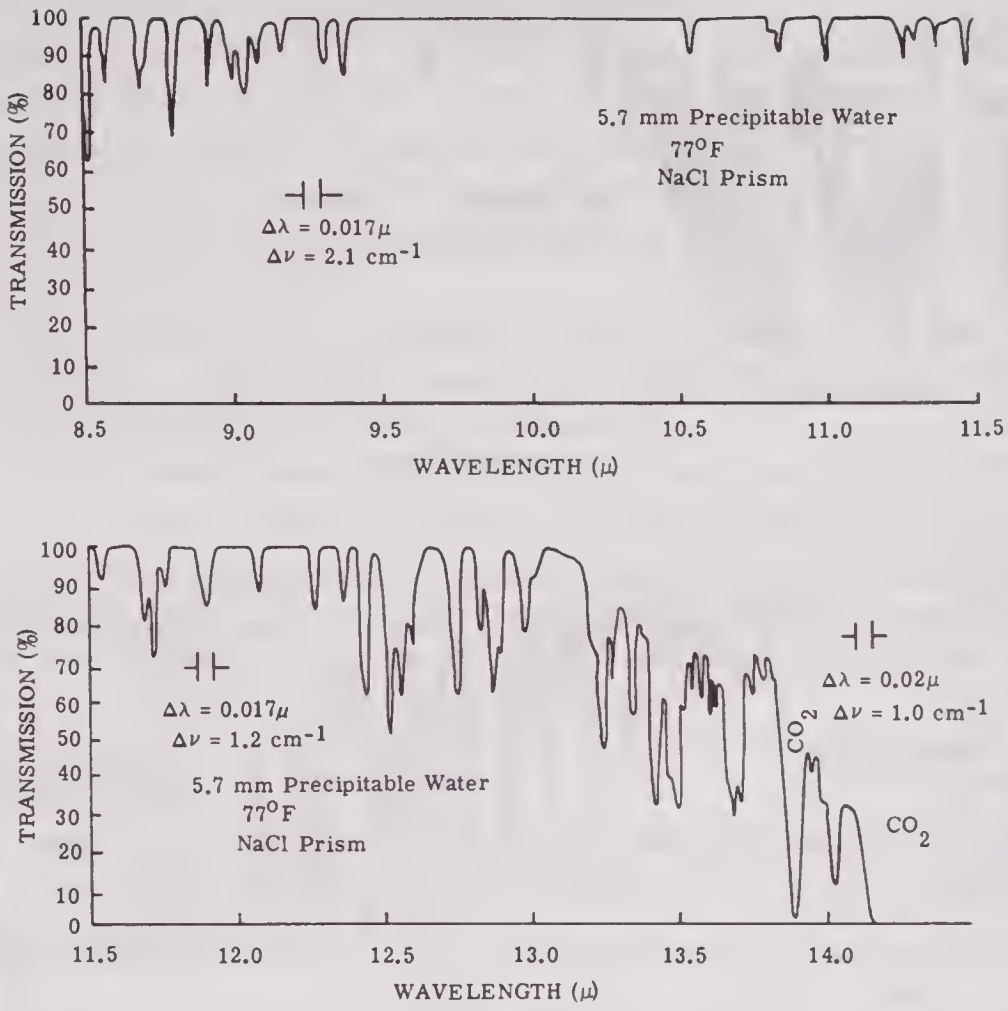


FIG. 6-114. Atmospheric transmission at sea level over a 0.3-km path, 8.5 to 14.0 μ [108].

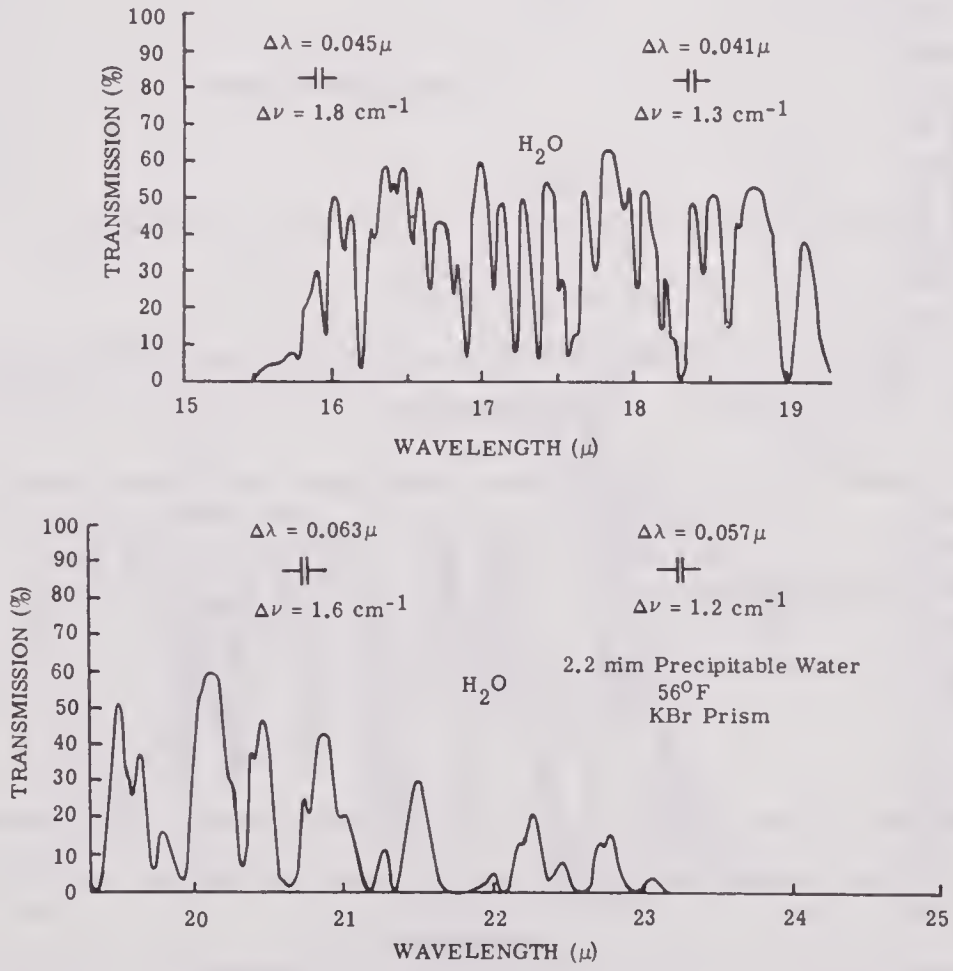


FIG. 6-115. Atmospheric transmission at sea level over a 0.3-km path, 15 to 25 μ [108].

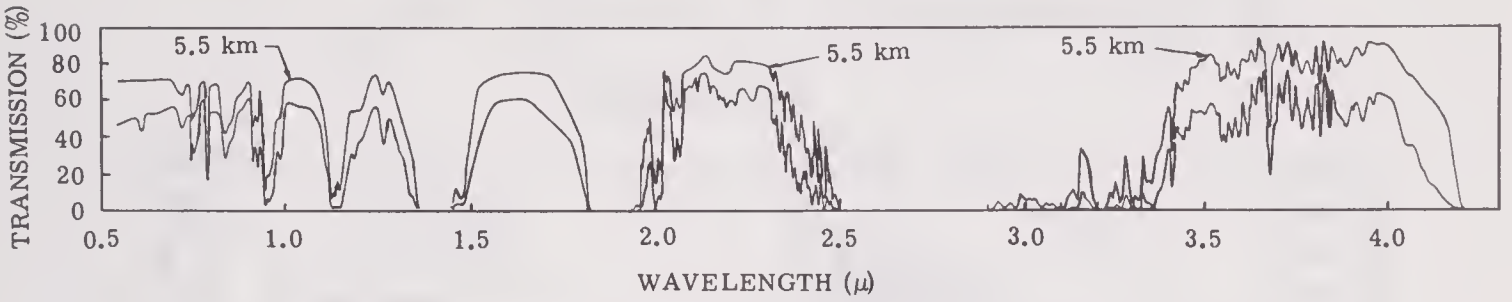


FIG. 6-116. Atmospheric transmission at sea level over 5.5- and 16.25-km paths, 0.5 to 4.0 μ [108].

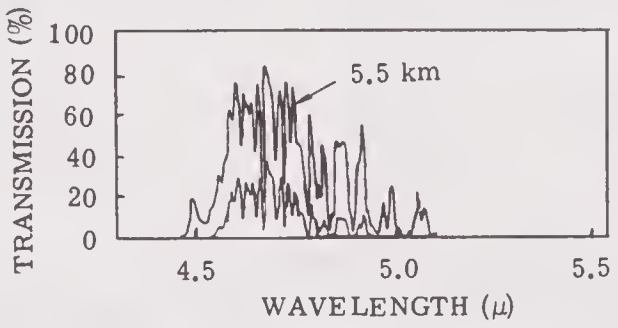
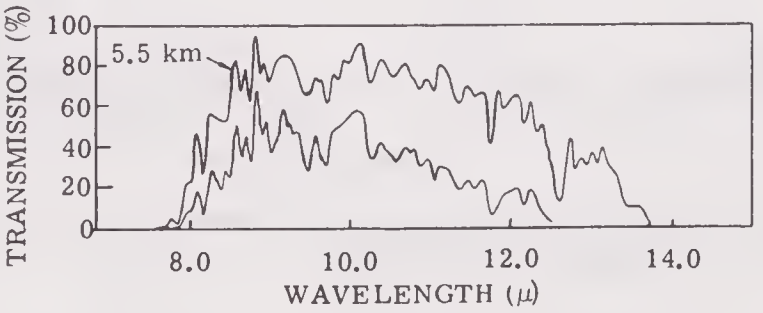


FIG. 6-117. Atmospheric transmission at sea level over 5.5- and 16.25-km paths, 4.5 to 5.5 μ [108].



	5.5 km	16.25 km
R.H. (%)	51	53
Temp. (^{O}F)	64	68
H_2O in path (cm)	4.18	15.1
Transmission at 0.55 μ (%)	70	43

FIG. 6-118. Atmospheric transmission at sea level over 5.5- and 16.25-km paths, 8.0 to 14 μ [108].

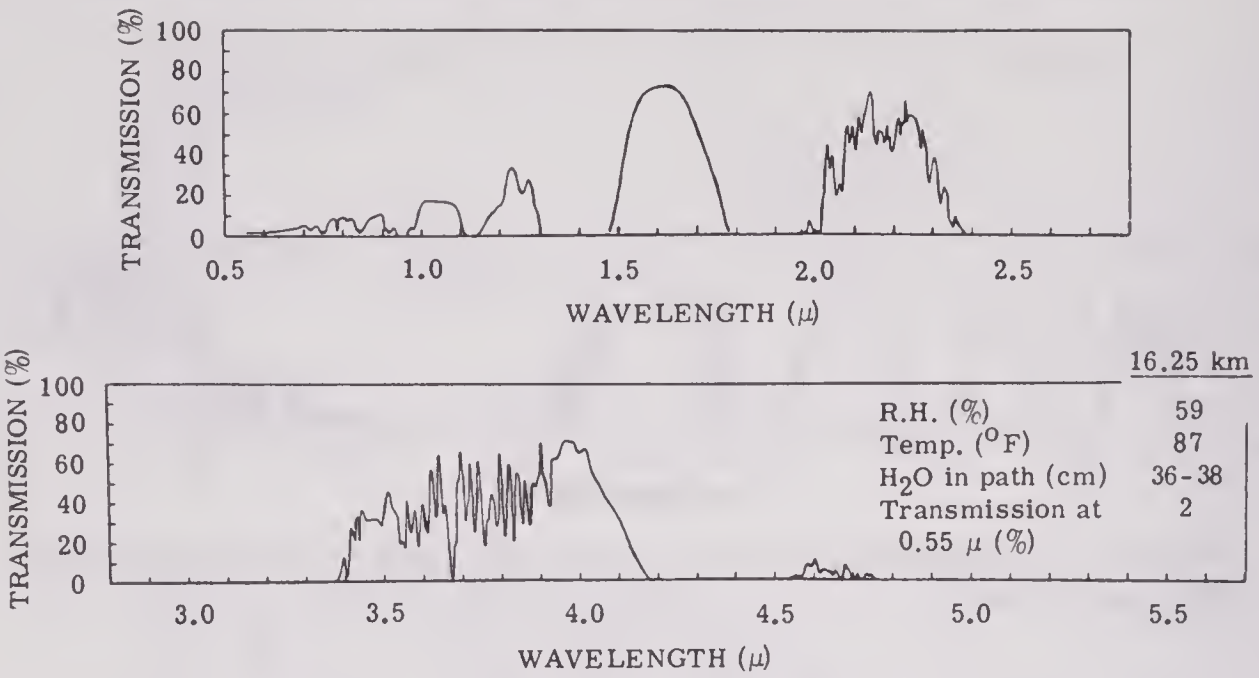


FIG. 6-119. Atmospheric transmission at sea level over a 16.25-km path [108].

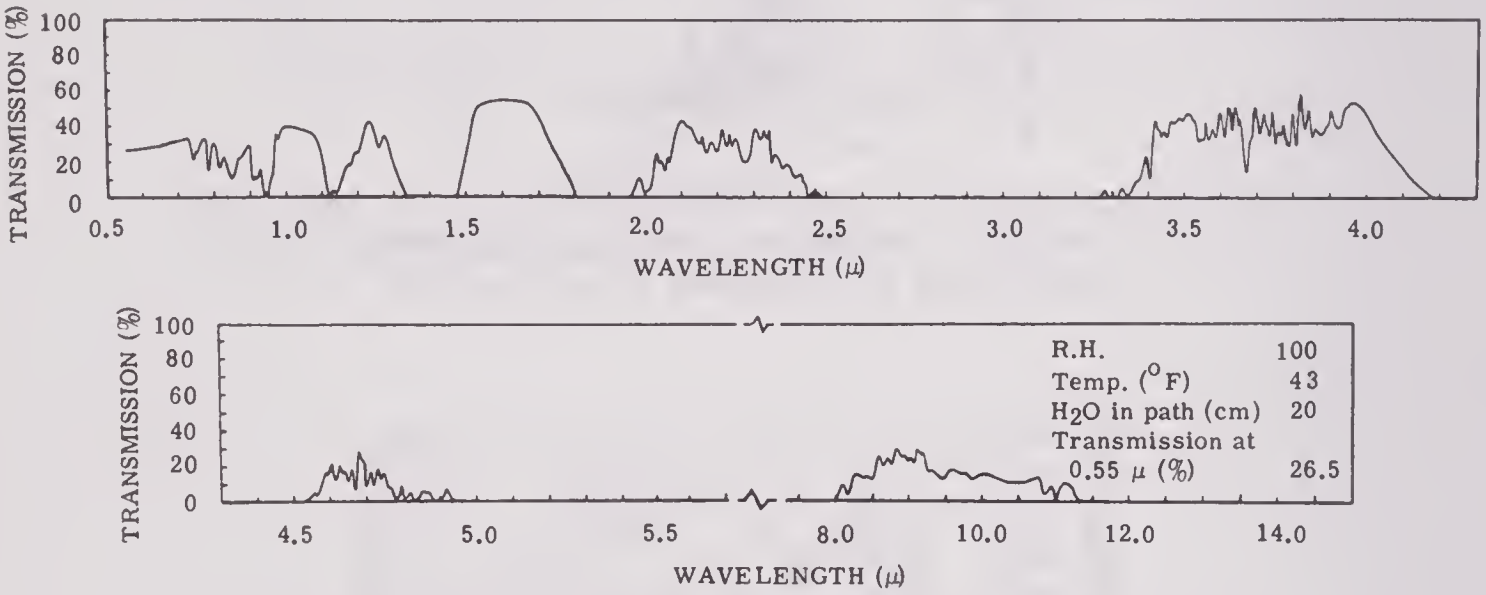


FIG. 6-120. Atmospheric transmission at 10,000 ft over a 27.7-km path when 0.55-μ transmission is 26.5% [108].

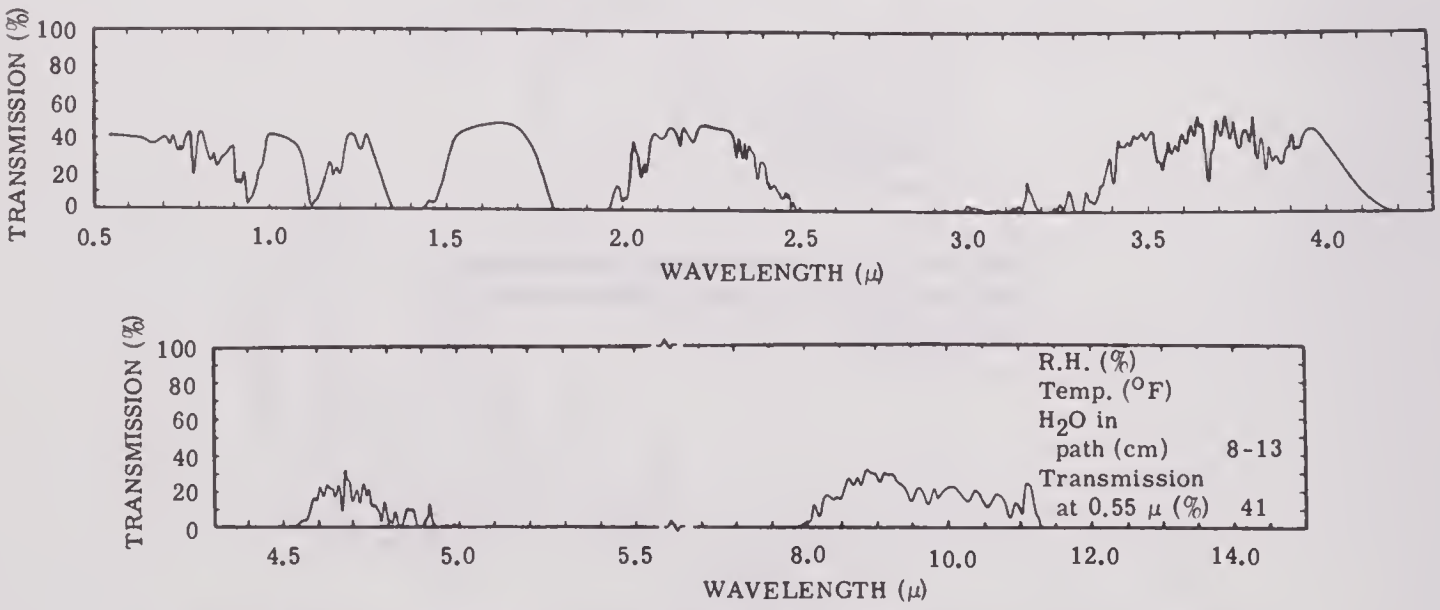


FIG. 6-121. Atmospheric transmission at 10,000 ft over a 27.7-km path when 0.55-μ transmission is 41% [108].

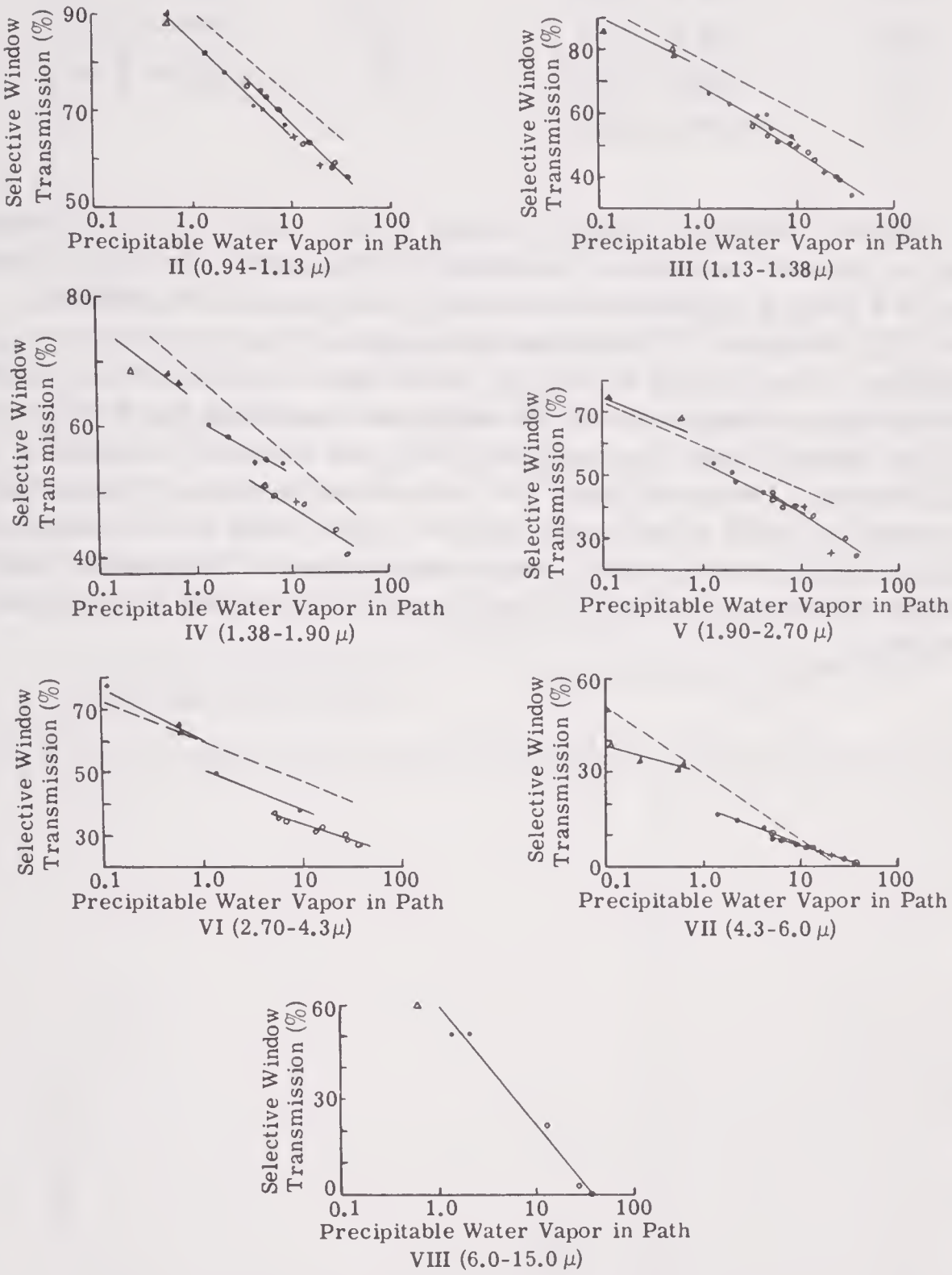


FIG. 6-122. Selective-window transmission [109].

TABLE 6-7. ATMOSPHERIC WINDOW DEFINITION [109].

Window No.	*Wavelength Limits (μ)*	*Window No.*	*Wavelength Limits (μ)*
I	0.72 to 0.94	VI	2.70 to 4.30
II	0.94 to 1.13	VII	4.30 to 6.00
III	1.13 to 1.38	VIII	6.00 to 15.00
IV	1.38 to 1.90	IX	15.00 to 25.00
V	1.90 to 2.70		

6.9.1.1. *Infrared Radiation Through Clouds* [110]. Clouds become increasingly transparent to infrared radiation as wavelength is increased. The optical density or "thickness" of a cloud is approximately inversely proportional to its visibility.

Figures 6-123 through 6-127 show cloud attenuation of infrared radiation as a function of visibility at an altitude of 2500 ft. These data were obtained by viewing an infrared source with a detector at various ranges and measuring the change in energy viewed by the detector when the measuring path was obscured by clouds of various measured densities. The optical density of a cloud was measured by determining the maximum range at which a "maximum contrast" object could be discerned in a given cloud. Figures 6-123 through 6-127 relate attenuation to "thickness of cloud," and the results do not necessarily apply to the various types of clouds encountered at different altitudes.

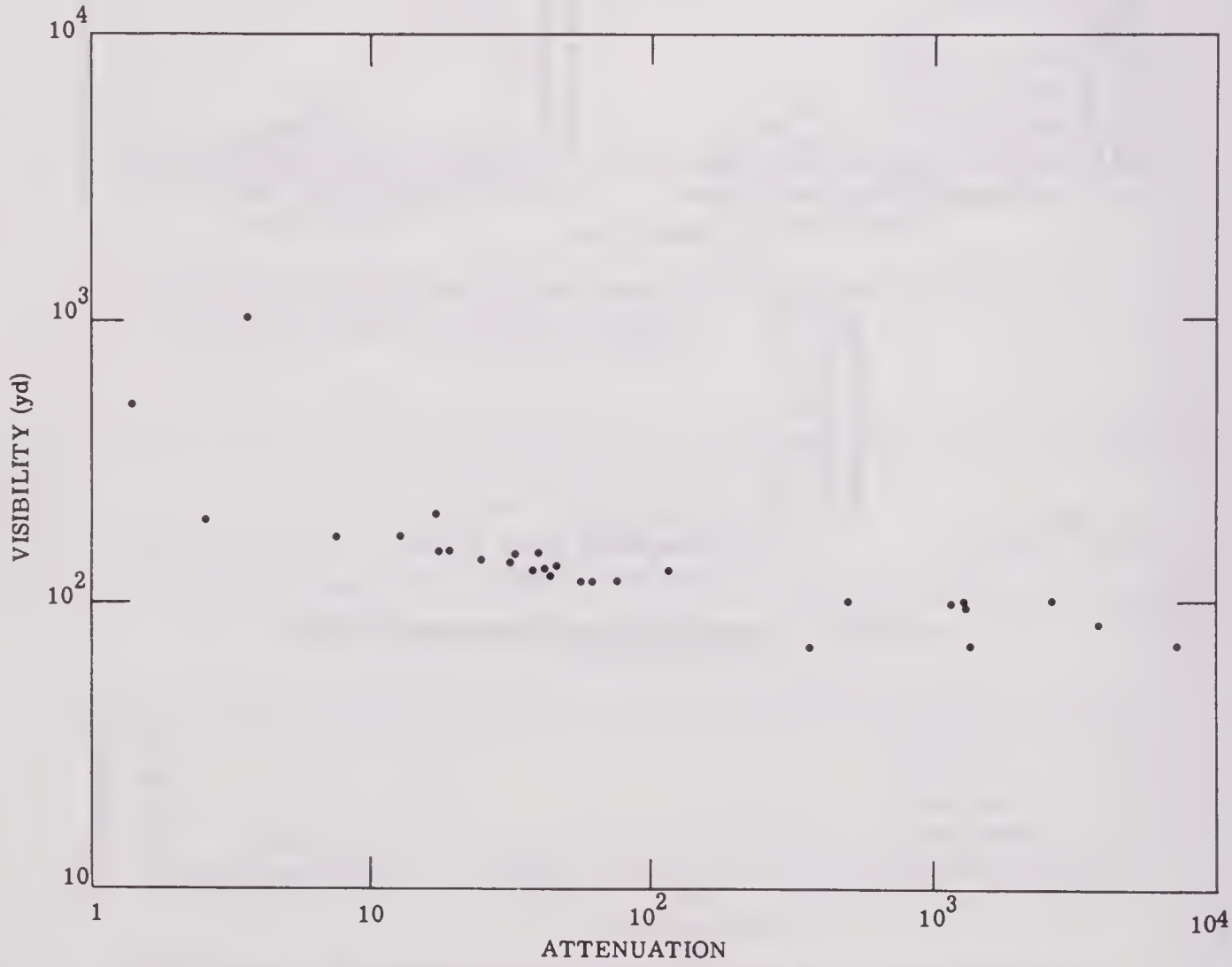


FIG. 6-123. Cloud attenuation *vs.* cloud visibility, 130 yd measuring range (PbS detector) [110].

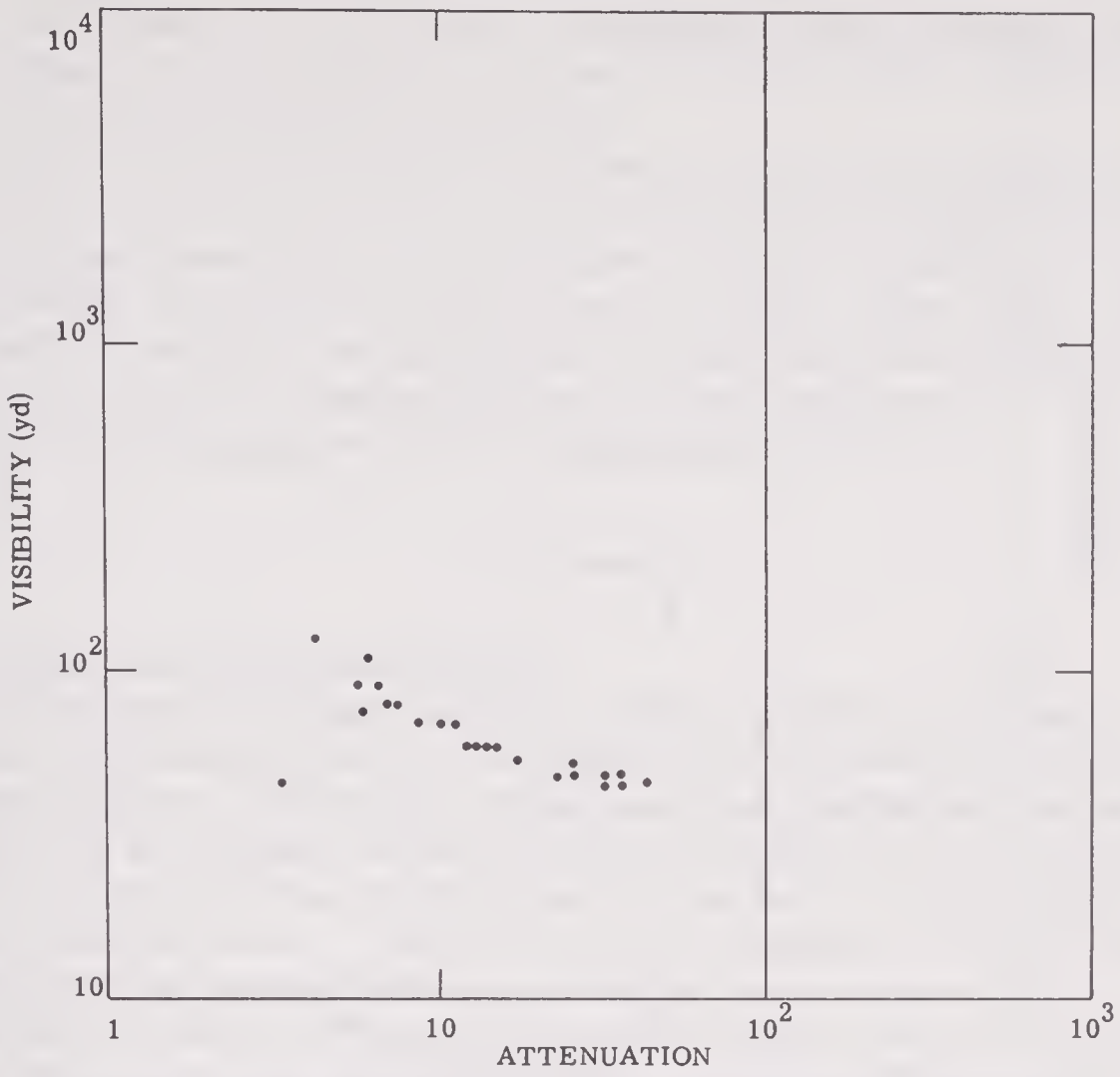


FIG. 6-124. Cloud attenuation *vs.* cloud visibility, 40 yd measuring range (PbS detector) [110].

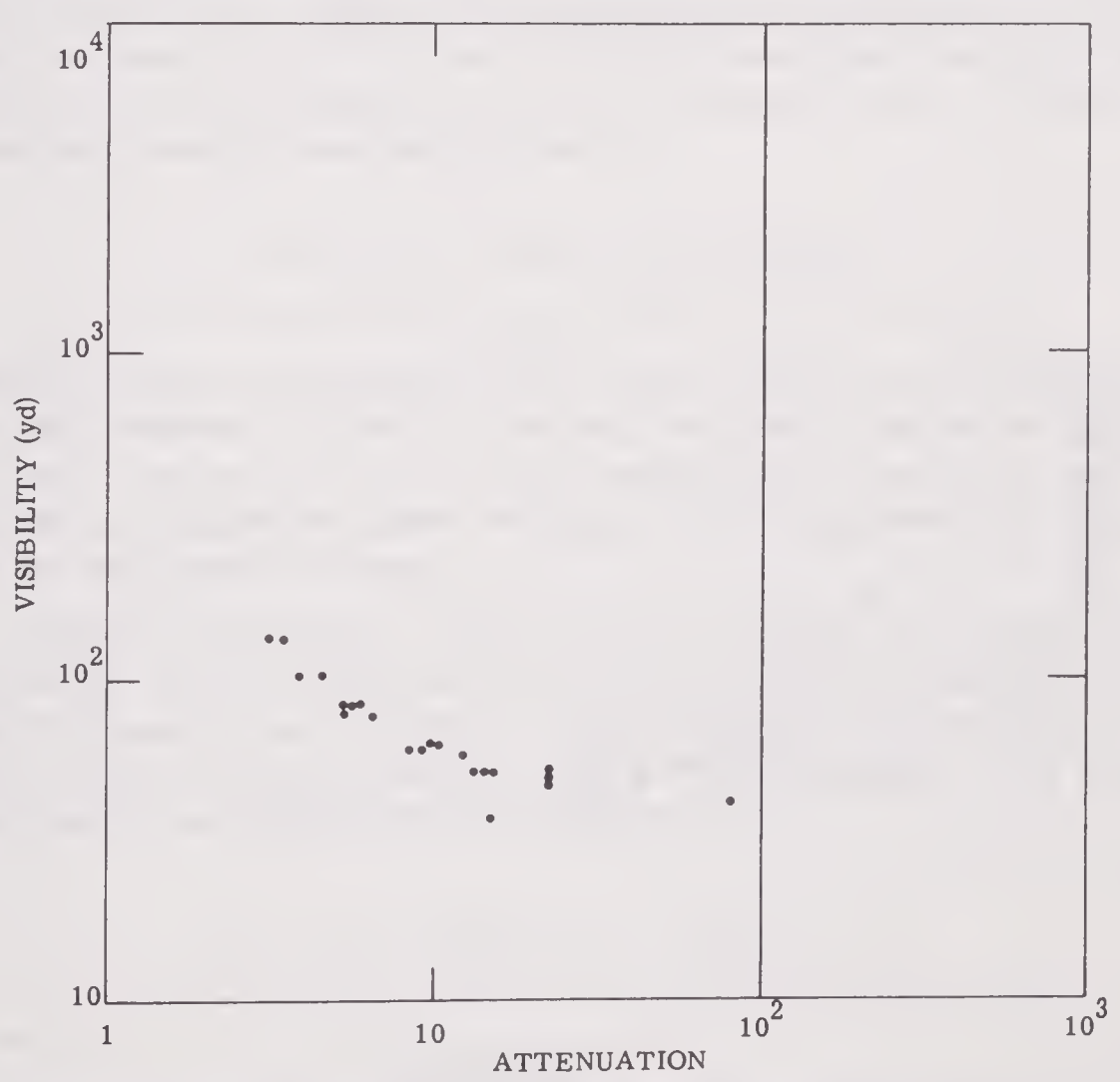


FIG. 6-125. Cloud attenuation *vs.* cloud visibility, 40 yd measuring range (PbSe detector) [110].

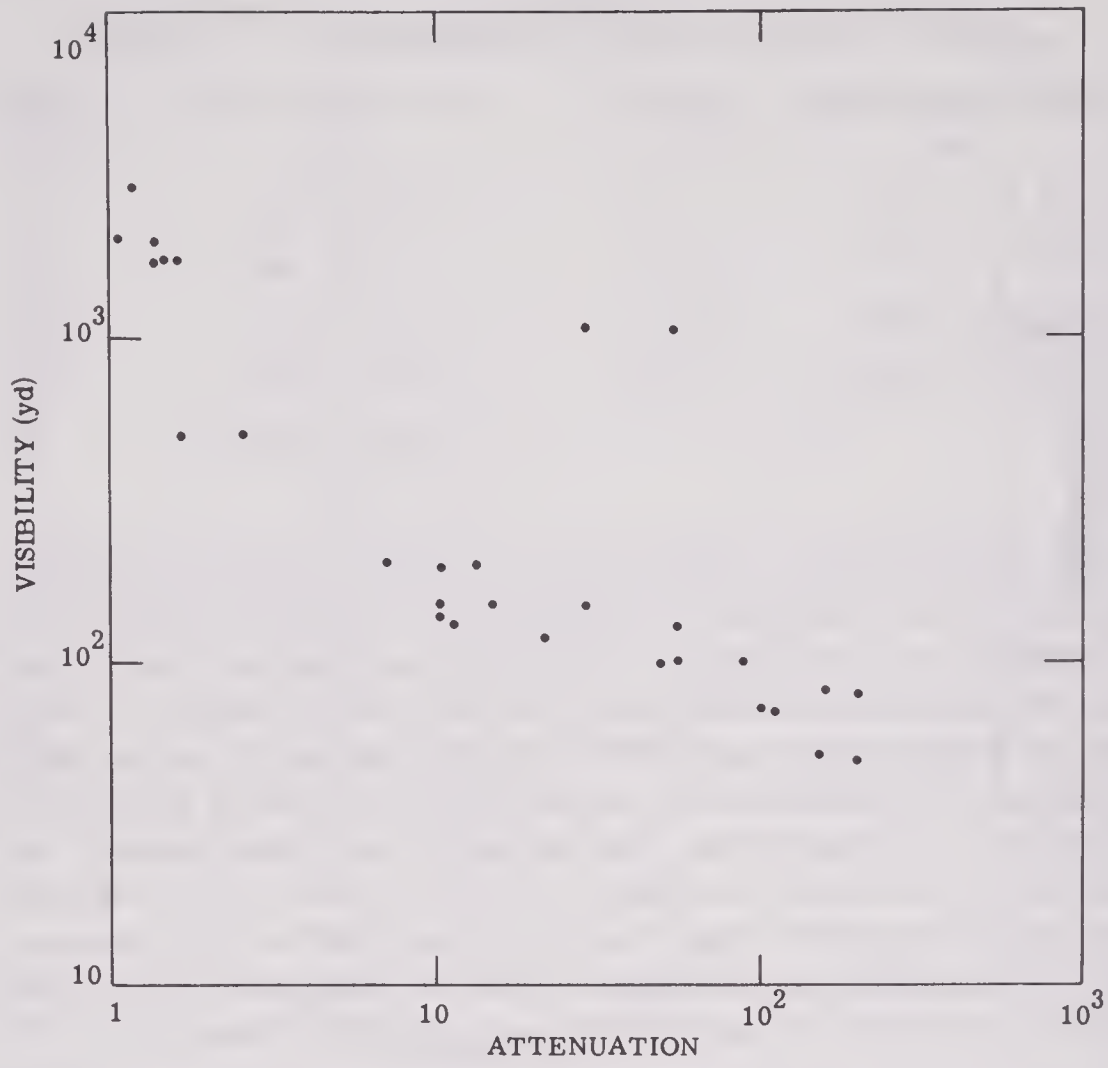


FIG. 6-126. Cloud attenuation *vs.* cloud visibility, 130 yd measuring range (PbTe detector) [110].

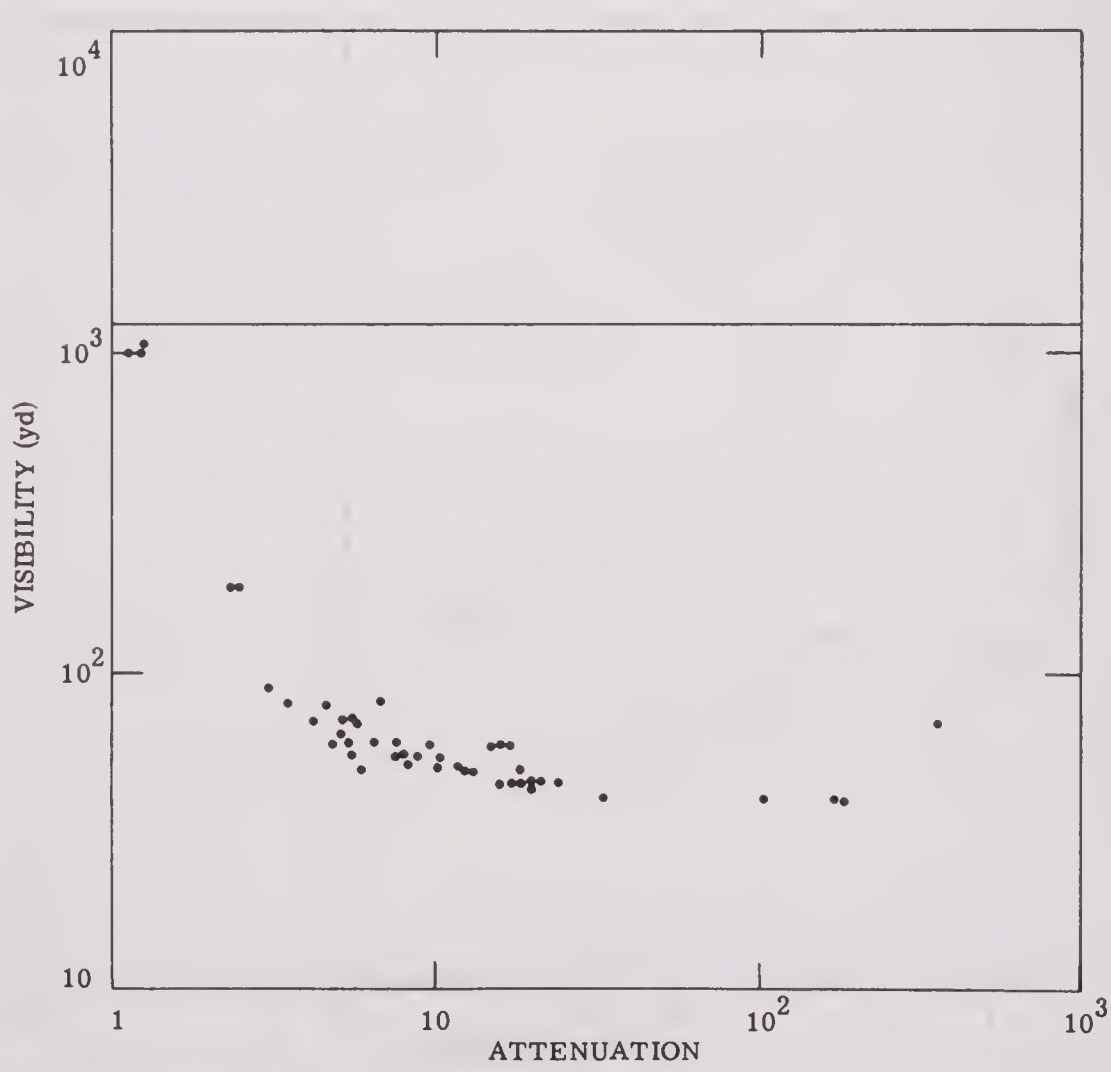


FIG. 6-127. Cloud attenuation *vs.* cloud visibility, 40 yd measuring range (PbTe detector) [110].

6.9.2. Slant Paths. Few quantitative data are available on infrared transmission over slant paths. The transmission must generally be predicted on the basis of various theoretical calculations discussed in this section.

Both pressure and temperature vary along most slant paths through the atmosphere. The half-width of the spectral lines varies with pressure according to Eq. (6-3). Both the half-width and intensity of the spectral lines vary with temperature. In addition, the concentration of the absorbing gases may vary along a slant path. Because of these factors it follows that the equations for slant path absorptance must necessarily be more complex than the corresponding homogeneous path equations.

The mass of absorbing gas per unit area along a slant path is

$$u = \int_0^L \rho_a \, d\ell \tag{6-54}$$

where ρ_a is the density of the absorbing gas

ℓ is the length of the path measured along the direction of the path.

Reference [111] describes elaborate calculations for the determination of the amount of gas along a slant path. From Fig. 6-128, an equivalent sea-level path can be obtained for any gas that is uniformly distributed in the atmosphere, such as CO_2, N_2O, CH_4, or CO. The quantity d_h is the equivalent path at sea level for a horizontal path 1 km long at the indicated altitude. The quantity d_v is the equivalent path at sea level for a vertical path from the indicated altitude to infinity.

It is necessary to assume a curve for the variation of H_2O concentration with height in order to obtain equivalent paths for this gas. In [111], a dry stratosphere with a constant mixing ratio in the stratosphere is assumed. There is no agreement at the present time between those who favor this distribution and those who favor a wet stratosphere (Sec. 6.1.5.2). Figure 6-129 gives equivalent centimeters of precipitable water vapor corrected to sea level for a 1-km horizontal path (w_h) at the indicated altitude and for a vertical path (w_v) from the indicated altitude to infinity. [111]

The absorptance along a slant path is given by

$$A\,\Delta\nu = \int_{\Delta\nu} \left[1 - \left(\exp - \int_0^u ku'\,du'\right)\right] d\nu \tag{6-55}$$

The coefficient is given by Eq. (6-2), (6-4), or (6-6). There are two methods for the solution of this problem. In the first method, analytical solutions are obtained for Eq. (6-53) under various conditions. In the second method, equivalent path lengths and pressures are obtained for a horizontal path so that its absorptance is the same as that along the slant path of interest.

6.9.2.1. *Analytical Solutions.* When the nonoverlapping line approximation is valid (Sec. 6.4.3), an exact expression can be obtained for absorptance over a slant path when the absorbing gas is uniformly mixed throughout the atmosphere and the temperature variation is small enough so that it can be neglected. The absorptance over a slant path at an angle θ with the vertical is [112]

$$A\,\Delta\nu = 2\pi\alpha_1\gamma \left(\frac{\alpha_0}{\alpha_1}\right)^{\gamma} \left[P_\gamma(z) - \left(\frac{\alpha_0}{\alpha_1}\right) P_{\gamma-1}(z)\right] \tag{6-56}$$

where $z = \dfrac{1}{2}\left(\dfrac{\alpha_0}{\alpha_1} + \dfrac{\alpha_1}{\alpha_0}\right)$ and $\gamma = \dfrac{Sc_0\, p_s \sec\theta}{2\pi\, g\, \alpha_s} = \dfrac{Su}{2\pi(\alpha_1 - \alpha_0)}$.

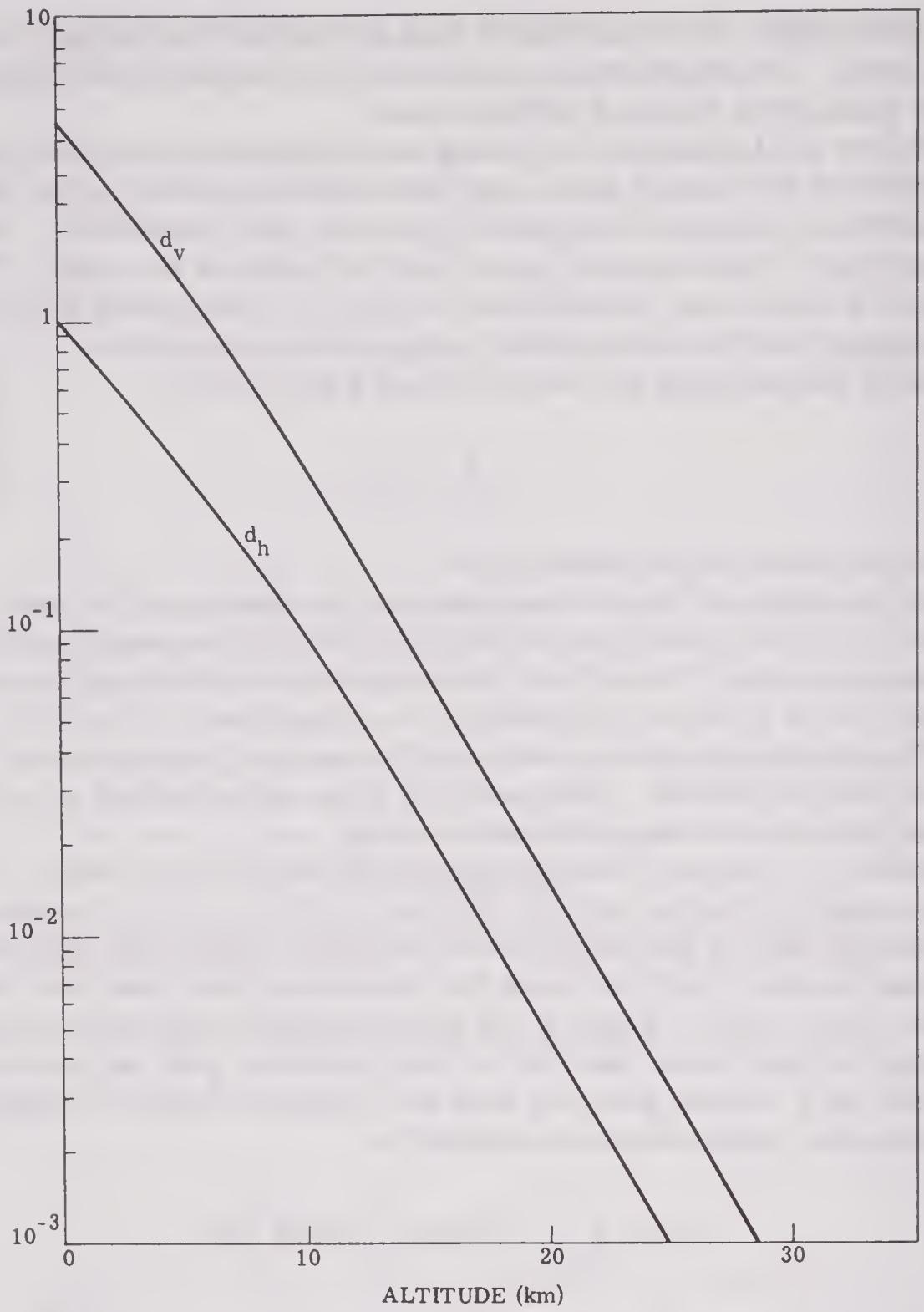


FIG. 6-128. Equivalent sea-level path length for any gas uniformly distributed in the atmosphere, *e.g.*, CO_2, N_2O, CH_4, CO [111].

α_0 and α_1 are the half-widths of the spectral lines at the beginning and end of the slant path, c_0 is the fractional concentration of the uniformly mixed absorbing gas, and $P_\gamma(z)$ is the Legendre function of order γ. Equation (6-56) is for a single line, but the absorptance values for any number of different lines can be added when they do not overlap. The parameter γ is similar for slant paths to x for homogeneous paths. The absorptance varies linearly with γ when $\gamma >> 1$ and as $\gamma^{1/2}$ when $\gamma >> 1$.

When γ is an integer, Eq. (6-56) can be written as a polynomial in z. For example, when $\gamma = 1$, the absorptance is

$$A\,\Delta\nu = \pi\alpha_1\left[1 - \left(\frac{\alpha_0}{\alpha_1}\right)^2\right] \tag{6-57}$$

An exact expression for the absorptance of an Elsasser band can be obtained [112] if the same assumptions are made as for the derivation of Eq. (6-56).

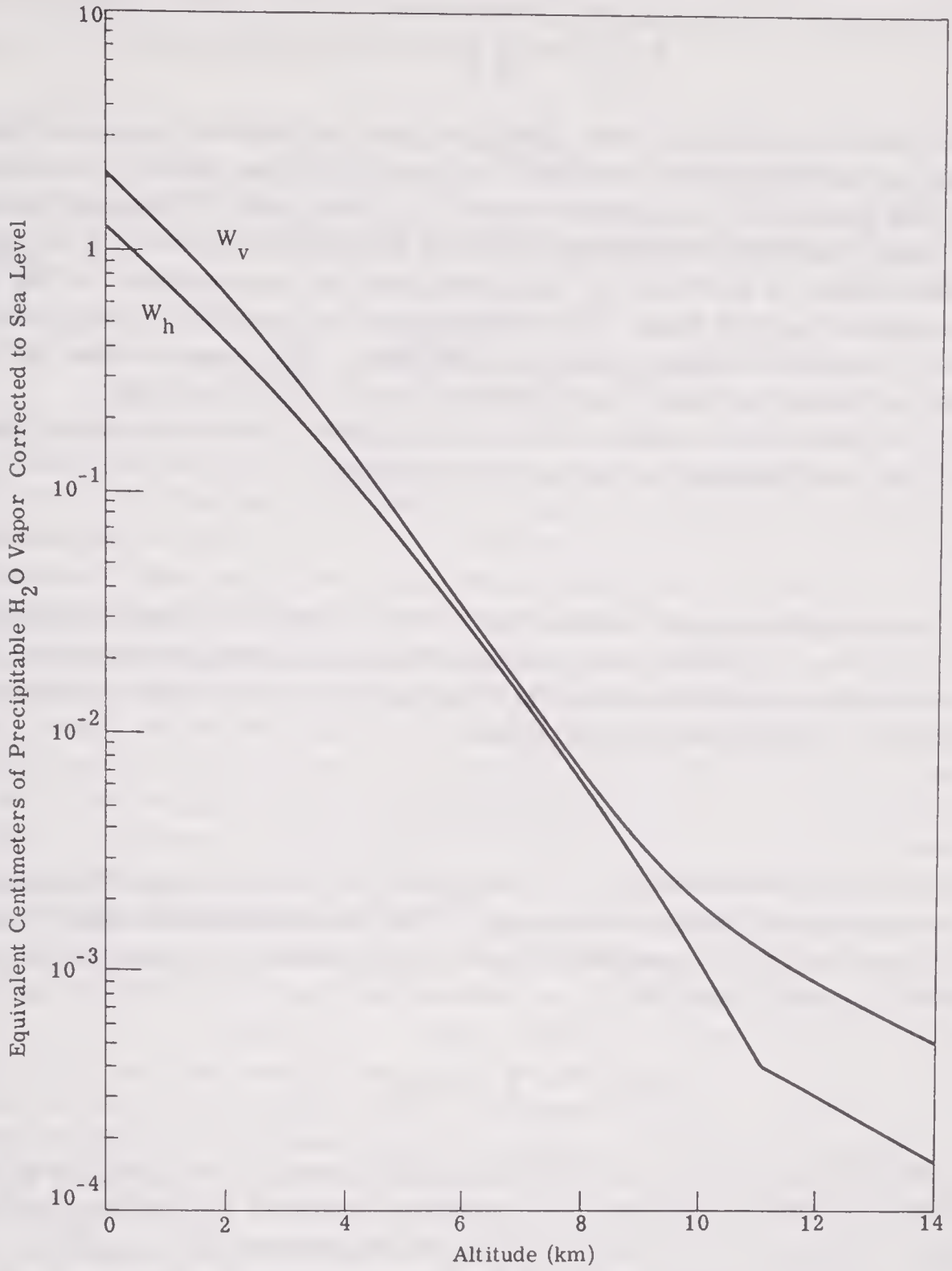


FIG. 6-129. Equivalent centimeters of precipitable water vapor [111].

The absorptance by a statistical band (Sec. 6.3.2) is still given by Eq. (6-21) or (6-22) when $A_{st,D}$ is interpreted to be the slant-path absorptance of a single spectral line over the frequency interval D. Thus, the absorptance by a statistical band can be obtained immediately when any expression for the absorptance of a single line is derived. The previous expressions for the random Elsasser (Sec. 6.3.3) and quasirandom (Sec. 6.3.4) models can be generalized in similar ways.

The changing mixing ratio of water vapor over long slant paths must be taken into account. It is still possible to obtain analytic expressions for the absorptance which converge rapidly in most cases of practical interest. This case is discussed in detail in [113] and [114].

6.9.2.2. *Correspondence Between Slant and Homogeneous Paths.* When the weak-line approximation is valid, a homogeneous path with an amount of absorbing gas u_h and at a temperature such that the individual line intensities are S_{ih}, has the same value of the absorptance as a slant path when [113, 114]

$$u_h \sum_{i=1}^{N} S_{in} = \sum_{i=1}^{N} \int_0^u S_i\,(u')\,du' \tag{6-58}$$

The integral in this equation is taken along the slant path with the amount of absorbing gas u taken as the independent variable. In general, the line intensity S_i varies along the slant path because the temperature varies along the path. This result is valid for any line shape regardless of whether the line is pressure broadened. It is also valid for any distribution of spectral line intensities and for any variation in the spacing between adjacent spectral lines. The temperature and absorber concentration may vary in any prescribed manner along the slant path. The pressure does not need to be specified as long as the weak-line approximation (Sec. 6.4.1) is valid.

When the temperature variation of the line intensities S_i can be neglected along the slant path, Eq. (6-58) reduces to the very simple form

$$u_h = u \tag{6-59}$$

When the strong-line approximation (Sec. 6.4.2) is valid, and when the temperature variation of the line intensities must be taken into account, a useful result can only be obtained when the temperature variation of all the spectral lines in the interval can be represented by a single function, $s(T)$, so that

$$S_i = S_{ih}\,{}_s\,(T) \tag{6-60}$$

Since the temperature variations in the atmosphere are not too large, this equation can usually be satisfied with sufficient accuracy. Then, the absorptance along a slant path is the same as that for a homogeneous path whose pressure, p_h, path length, u_h, and temperature T_h, satisfy [113, 114]

$$p_h u_h T_h^{-1/2} = \int_0^u s(T) p\; T^{-1/2}\, du' \tag{6-61}$$

The only assumptions made in the derivation of this equation are that the strong-line approximation is valid, that there is a pressure-broadened half-width, and that the temperature variation of the line intensities can be represented by Eq. (6-60). Equation (6-61) is valid for any variation of the line intensities, half-widths, and spacing between the lines within the spectral interval.

If, in addition, the temperature variation along the slant path can be neglected and this temperature is the same as that along the homogeneous path, then $s(T) = 1$ and Eq. (6-61) simplifies to [113, 114]

$$p_h u_h = \int_0^u p\,du' \tag{6-62}$$

If the absorbing gas is also distributed uniformly along the slant path, and if there is the same amount of absorbing gas along the homogeneous and slant paths, Eq. (6-62) reduces to [113, 114]

$$p_h = \frac{1}{2}\,(p_0 + p_1) \tag{6-63}$$

where p_0 and p_1 are the values of the pressure at the two ends of the slant path.

If the variation with height of the fractional concentration of a gas such as water vapor can be written as some power of the pressure

$$c = c_0 \left(\frac{p}{p_0}\right)^{\ell} \tag{6-64}$$

where c_0 is the fractional concentration when the pressure is p_0 and ℓ is any number, then [113]

$$u_h p_h = \frac{c_0 \sec \theta}{(\ell + 2) g {p_0}^{\ell}} ({p_1}^{\ell+2} - {p_0}^{\ell+2}) \tag{6-65}$$

where g is the acceleration of gravity.

This method of obtaining the absorption along a slant path from homogeneous path data is discussed in detail in [113, 114]. Reference [115] gives slant-path absorption from laboratory data, using the equations as a basis for calculations. Appropriate equivalent sea-level paths for various slant paths were first calculated, and then applied to laboratory absorption measurements.

An important approximate expression, which is valid over a wide range of pressure and path length, can be obtained for slant-path absorption by combining the weak-line and strong-line methods of determining slant-path absorption from homogeneous path data. When the weak-line approximation is valid, a value of u_h is determined from Eq. (6-58), which determines the appropriate path length for the homogeneous path. On the other hand, a value of the product $p_u u_h$ is determined from Eq. (6-61), (6-62), (6-63), and (6-65), when the strong-line approximation is valid.

It is possible to satisfy simultaneously any pair of these weak- and strong-line equations, since the weak-line equation determines the value of u_h and the strong-line equation can then be solved for p_h. These results may then be substituted into any equation which has been derived, or used in conjunction with any table of measured or calculated values of absorptance along a homogeneous path. The slant-path absorptance calculated in this manner necessarily agrees with the correct result in both the weak- and strong-line limits. In the intermediate region where both the weak- and strong-line approximations may be somewhat in error, the result derived in this manner provides a smooth interpolation between the absorptance curves that are valid in the weak- and strong-line limits. In general, the value obtained from this interpolation method is very close to the actual value for the slant-path absorptance.

As an example of this method, when Eq. (6-58) and (6-61) are combined it is found that

$$u_h = \frac{\sum_{i=1}^{N} \int_0^{\mu} S_i \, du'}{\sum_{i=1}^{N} S_{ih}} \tag{6-66}$$

and

$$p_h = \frac{{T_h}^{1/2} \sum_{i=1}^{N} \int_0^{\mu} S_i p T^{-1/2} \, du'}{\sum_{i=1}^{N} \int_0^{u} S_i \, du'} \tag{6-67}$$

When these values are substituted into any expression for the absorptance along a homogeneous path, the resulting equation for the absorptance is necessarily correct for a slant path in both the weak- and strong-line limits.

A more complete discussion of possible combinations of weak- and strong-line expressions and their use with various equations for the absorptance along a homogeneous path is given in [113]. This method is a generalization of the interpolation procedure described in [116] and [117].

A number of calculations of slant-path absorptances over various regions of the spectrum have been made. Those first described in [115] are discussed earlier in this section.

Reference [118] gives the transmissivity along a vertical path from a given altitude to the top of the atmosphere. The data are based on the equivalent sea-level paths given in reference [115]. The calculations apply only to CO_2 and H_2O absorption from 2 to 5 μ.

Reference [111] contains graphs of equivalent sea-level paths for the various absorbing gases, vapors, and haze which may be encountered along an atmospheric path. The absorptivity calculations are based on laboratory data and assume the 1959 ARDC model atmosphere. However, methods for correcting the results for various other conditions are given. The absorptivity of all important gases is considered from 1 to 20 μ and the H_2O data is extended out to 4 mm.

Reference [119] gives slant-path absorptivities for H_2O and CO_2 from 1 to 10.8 μ. Tables are given for computing the CO_2 and H_2O amounts along the line of sight and the equivalent pressures to be used for the absorptivity calculations. Attempts are made to fit each wavelength interval by the absorption model which represents best the data for that interval. Reference [120], from the transmission models for H_2O and CO_2 given in [119], presents simplified methods for determining equivalent paths. The effect of a curved earth is taken into account. The results are expressed in terms of three constants for CO_2 and one constant for H_2O, which are tabulated at various wavelengths.

A study of the absorption in the 2.7-μ region of H_2O and CO_2, with simplified band models, is described in [121]. An empirical expression to fit the Elsasser band model to the 4.3-μ CO_2 band is discussed in [122]. The slant-path transmission is calculated by dividing the atmosphere into small layers and applying Lambert's law to each layer starting from the top of the atmosphere.

A detailed calculation of the transmissivity from the top of the atmosphere down to certain altitudes and at a number of angles to the vertical is given in [123]. Equations similar to (6-66) and (6-67) have been used to correlate the appropriate quantities for slant and homogeneous paths. The calculations are based on the transmissivity tables described in [56] and [57]. The transmissivity is given for each 5-cm^{-1} interval from 500 to 10,000 cm^{-1} for CO_2 and from 1000 to 10,000 cm^{-1} for H_2O.

6.10. Calculation Procedures

Previous sections of this chapter have discussed in detail the different aspects of the earth's atmosphere as related to atmospheric transmission. In order to determine the transmissivity for a specific slant path one must first determine the equivalent sea-level path for the absorber of interest. The second step is to find an absorption coefficient which is characteristic of the wavelength or the wavelength interval under consideration. This can be carried out by using the figures presented in this chapter or by measured coefficients given by Howard, Burch, and Williams [107]. The final step in determining path transmissivity is to select the functional relation between transmissivity and the product of the absorption coefficient and the reduced optical path which best suits the slant path under consideration.

The most tedious part of this process is the determination of the equivalent sea-level path. In general the integral relation

$$u_e = \int_{x_1}^{x_2} q(x) \left(\frac{P}{P_0}\right)^n \left(\frac{T}{T_0}\right)^m \rho(x)\, dx \tag{6-68}$$

must be evaluated along the atmospheric path for the specific set of atmospheric conditions. Here $q(x)$ is the mixing ratio along the slant path; p is the pressure along the slant path; $T(x)$ is the temperature along the slant path; $\rho(x)$ is the density of air along the slant path; m and n refer to the power of the temperature and pressure correction for line half-width. For high accuracy this relation is best evaluated by digital computer; however, Carpenter and Altschuler have each developed graphical techniques which apply with reasonable accuracy for a standard model atmosphere. Carpenter's procedure will be given here.

Adopting Carpenter's notation, h_s is the source altitude, h_d is the detector altitude, x_g is the ground range between source and detector, and h_0 is the minimum slant-path altitude. It is evident, as shown in Fig. 6-130, that the atmospheric paths can be separated into two types, A and B. Suppose the slant path were extended indefinitely in

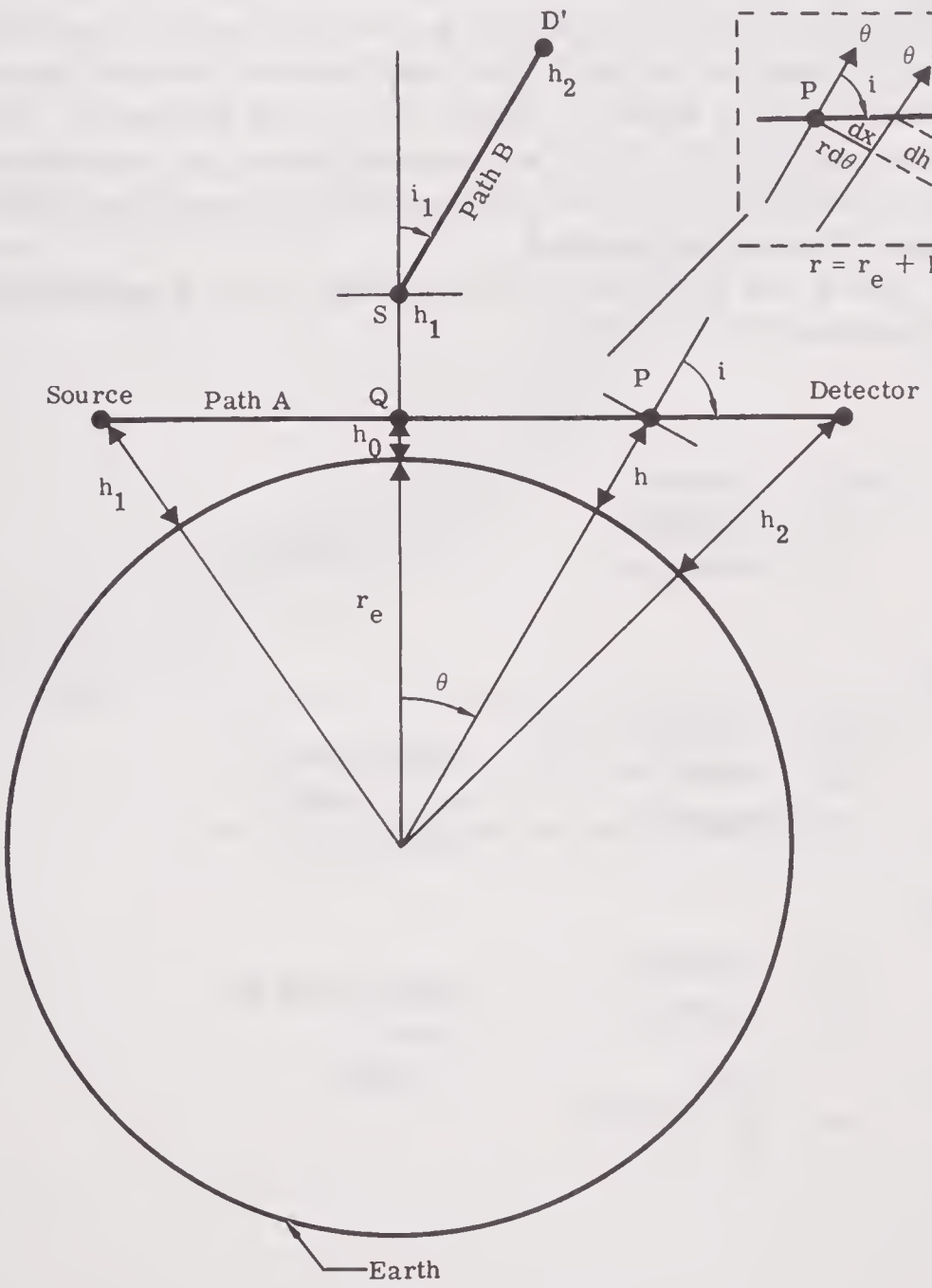


FIG. 6-130. Optical ray paths.

both directions. If it then intersected the earth, it would be of type B; otherwise it might be considered to be of type A, which is horizontal.

STEP 1

Usually h_s, h_d, and x_g are the known geometric parameters of the slant path. The first step is to determine for a given slant path whether the slant path is of type A or type B. Figure 6-131 is used for this purpose, and a trial-and-error procedure should be carried out.

To determine whether a path is of type A (the horizontal class) a series of very long horizontal paths has been constructed (Fig. 6-131). Such paths might be referred to as fundamental horizontal paths. Any slant path of type A will be a segment of one of the illustrated paths. Note that each fundamental path is uniquely specified by its minimum altitude h_0. Therefore if one wishes to determine whether a given slant path is of type A or B, two horizontal lines would be drawn on Fig. 6-131 with the first at the altitude of the source and the second at the altitude of the receiver. These horizontal lines will intersect many of the fundamental paths. If one of the intersected paths has approximately the appropriate ground distance, then this fundamental path contains as a segment the slant path of interest. From the intersection of the horizontal lines with the fundamental path the specific elevation angles at the source point and receiver point can be read together with the corresponding value of h_0 for the path.

From Fig. 6-130 it can be seen that point Q may or may not be a part of the desired slant path. If point Q does not fall on the slant path but the path is of type A, the path-determination procedure is as described above. If point Q falls within the slant path then it should be realized that Fig. 6-131 is symmetric about the ordinate axis. Under this condition the figure might be redrawn to include the unpublished half. The path determination would proceed as described.

Slant paths of type B can be divided into two cases. Case I applies whenever $86° > i > 90°$; case II applies if $0 > i > 86°$.

Example 1

$$\begin{aligned} \text{Let } h_d &= 15{,}000 \text{ ft} \\ h_s &= 30{,}000 \text{ ft} \\ x_g &= 150{,}000 \text{ ft} \end{aligned} \quad \rightarrow \quad \begin{aligned} &\text{type } A \\ &h_0 = 10{,}000 \text{ ft} \end{aligned}$$

Example 2

$$\begin{aligned} h_d &= 15{,}000 \text{ ft} \\ h_s &= 30{,}000 \text{ ft} \\ x_g &= 500{,}000 \text{ ft} \end{aligned} \quad \rightarrow \quad \begin{aligned} &\text{type } B, \text{ case I} \\ &h_0 = 10{,}000 \text{ ft} \\ &i = 1.10 \end{aligned}$$

Example 3

$$\begin{aligned} h_d &= 15{,}000 \text{ ft} \\ h_s &= 30{,}000 \text{ ft} \\ x_g &= \frac{\sqrt{3}}{2} 30{,}000 \text{ ft} \end{aligned} \quad \rightarrow \quad \begin{aligned} &\text{type } B, \text{ case II} \\ &l_0 = 0 \\ &i = 600 \end{aligned}$$

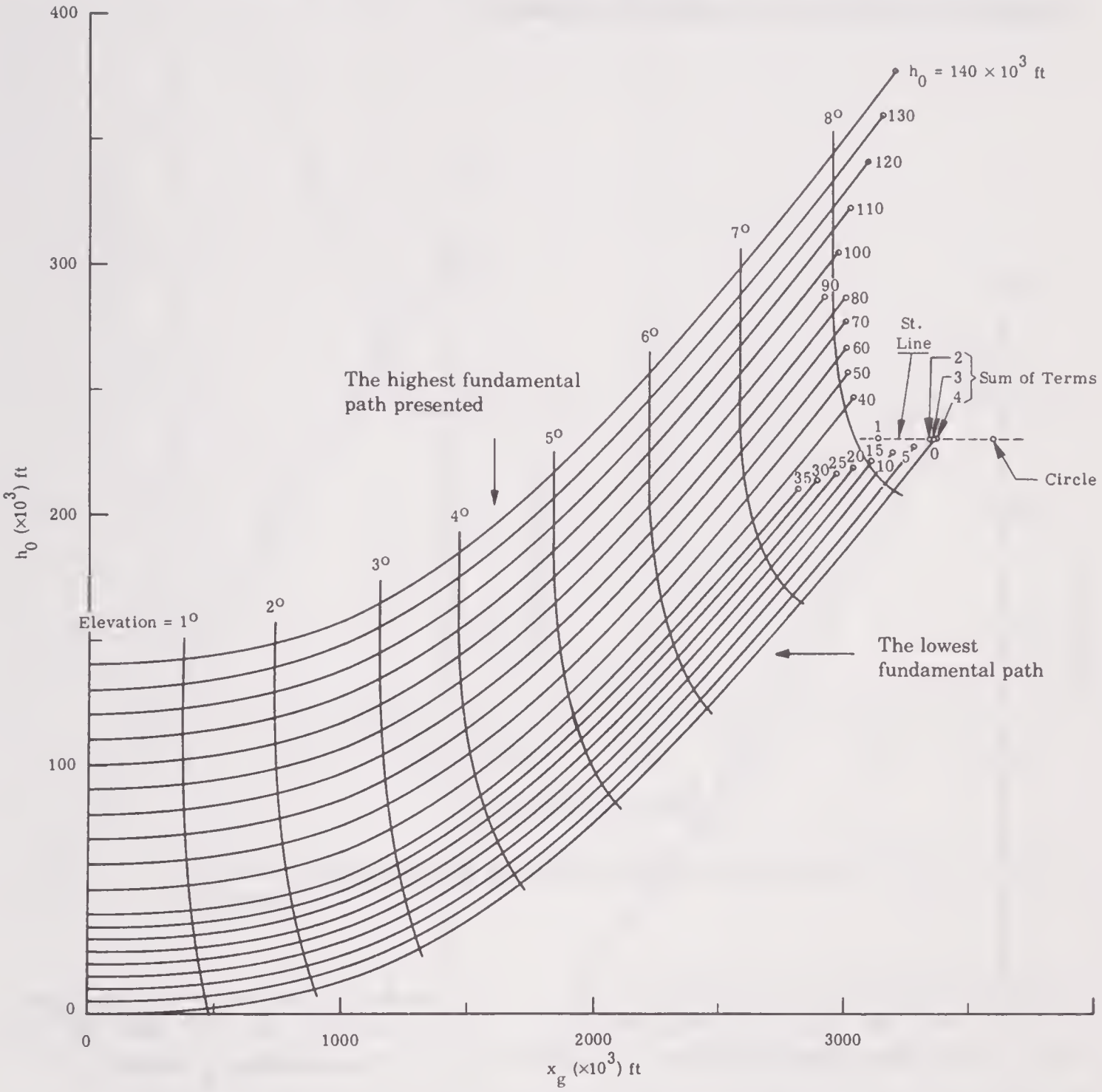


FIG. 6-131. Horizontal optical trajectories.

STEP 2

The second step for carbon dioxide is to determine the equivalent optical path for an infinite horizontal slant path, u_h, or a vertical path, u_v.

Use Fig. 6-132 to get u_h for atmospheric paths of type *A*

Use Fig. 6-132 to get u_h for slant paths of type *B*, case I

Use Fig. 6-133 to get u_v for slant paths of type *B*, case II, where in Fig. 6-133 the abscissa is either h_d or h_s, whichever is smaller.

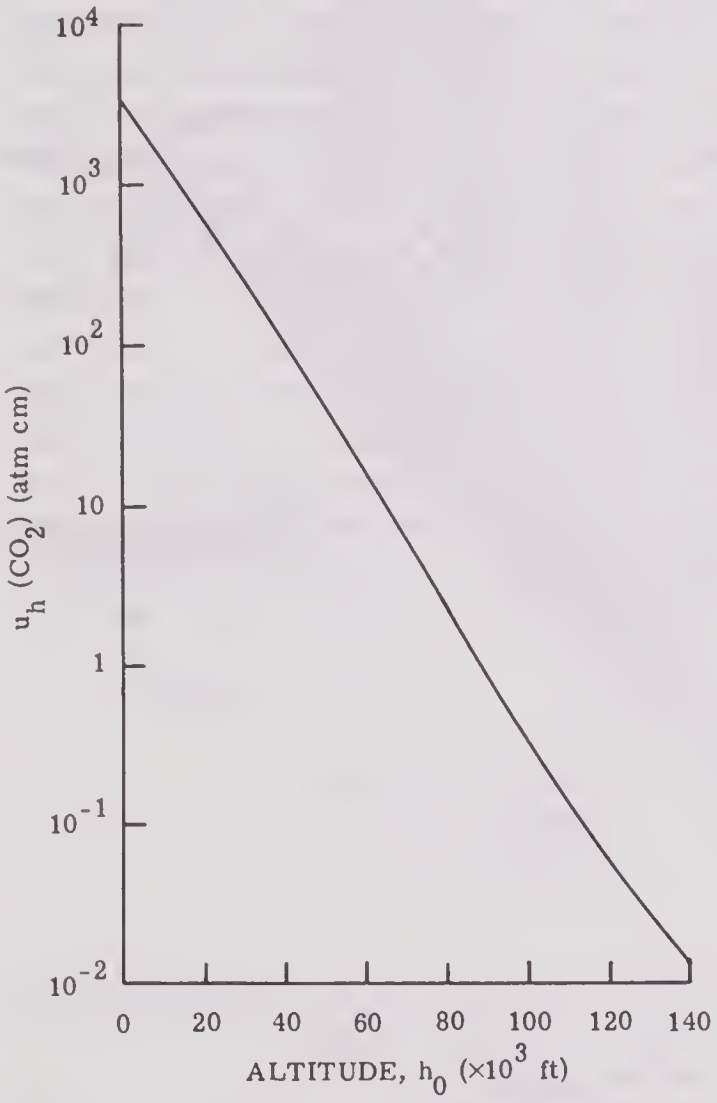


FIG. 6-132. Total horizontal equivalent optical path length for CO_2 *vs.* altitude.

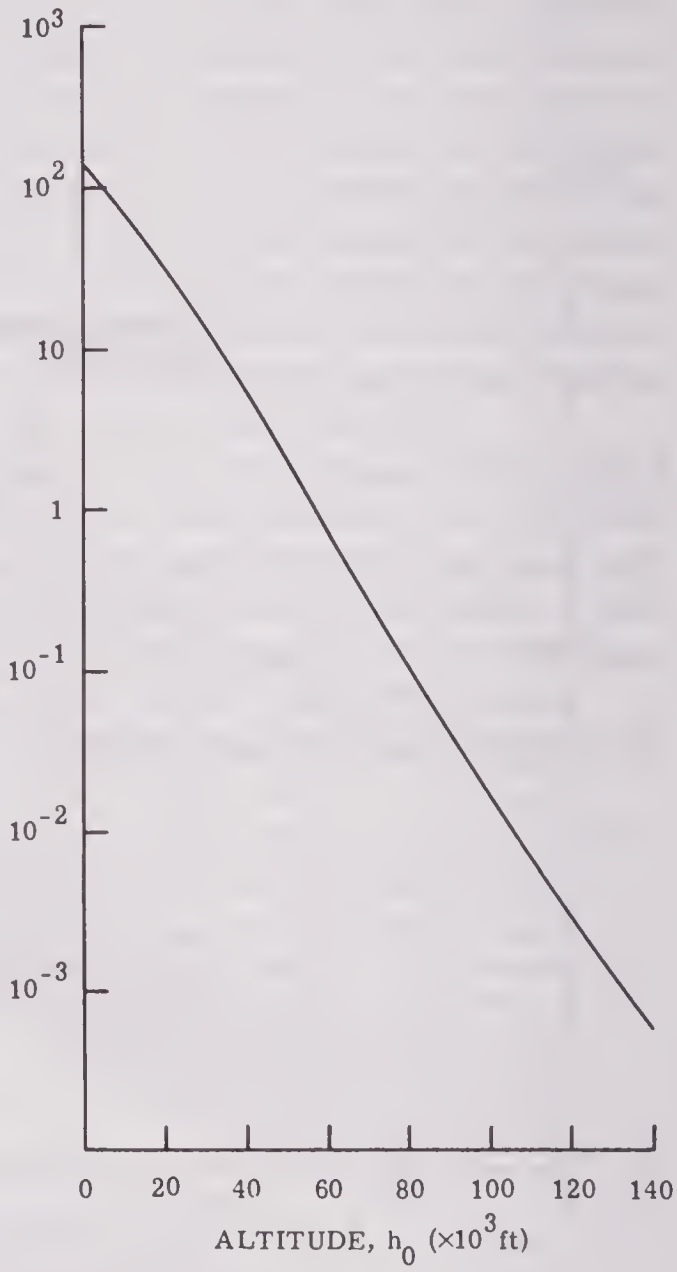


FIG. 6-133. Scale height for $P^2/T^{3/2}$

STEP 3

For Carpenter's procedure, step 3 is the determination of a scale height H_0'' which can be found in Fig. 6-134 for all atmospheric paths involving CO_2.

From steps 1 or 2 the value of h_0 has been determined. To find H_0'', one simply reads the appropriate value from Fig. 6-134.

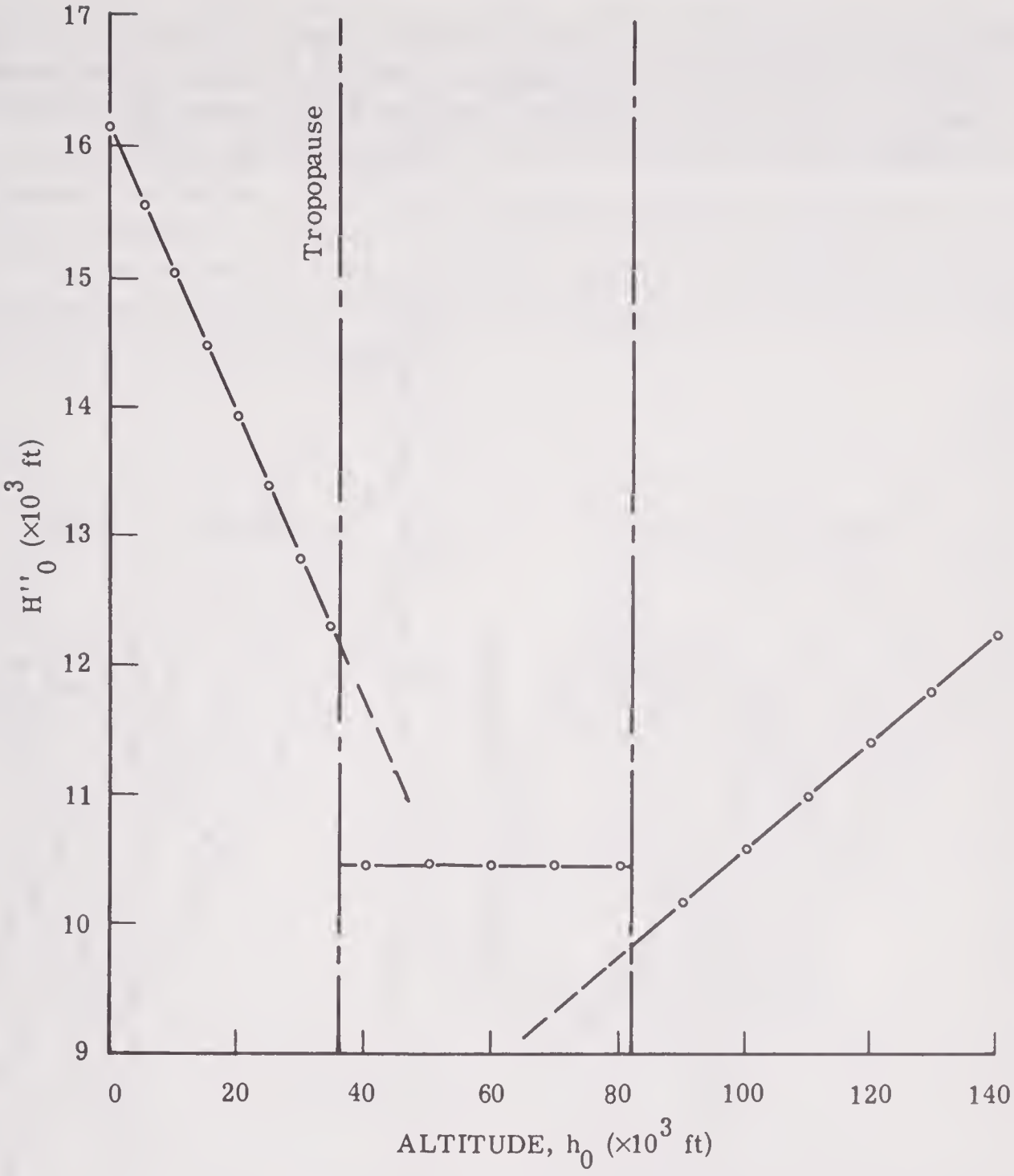


FIG. 6-134. Total vertical equivalent optical path length for CO_2 *vs.* altitude.

STEP 4

To compute the equivalent sea-level paths for atmospheric paths of type *A*, use

$$u_e = u_h \left[\operatorname{erf} \left(\frac{h_d - h_0}{H_0''} \right)^{1/2} + \operatorname{erf} \left(\frac{h_s - h_0}{H_0''} \right)^{1/2} \right] \tag{6-69}$$

For atmospheric paths of type *B*, case I, use

$$u_e = u_h \left[\operatorname{erf} \left(\frac{h_d - h_0}{H_0''} \right)^{1/2} \operatorname{erf} \left(\frac{h_s - h_0}{H_0''} \right)^{1/2} \right] \tag{6-70}$$

For atmospheric paths of type *B*, case II, use

$$u_e = \sec i \; u_v \, (h_d) \left[1 - \exp - \left[\frac{h_s - h_d}{H_0''} \right] \right] \tag{6-71}$$

STEPS 2-4, *for water vapor when the mixing ratio is known.* When the mixing ratio m is constant for all altitudes, the procedure is identical for both carbon dioxide and water vapor with the following exceptions: use Fig. 6-135 instead of Fig. 6-132, use Fig. 6-136 instead of Fig. 6-133 and use Fig. 6-135 again for Fig. 6-135.

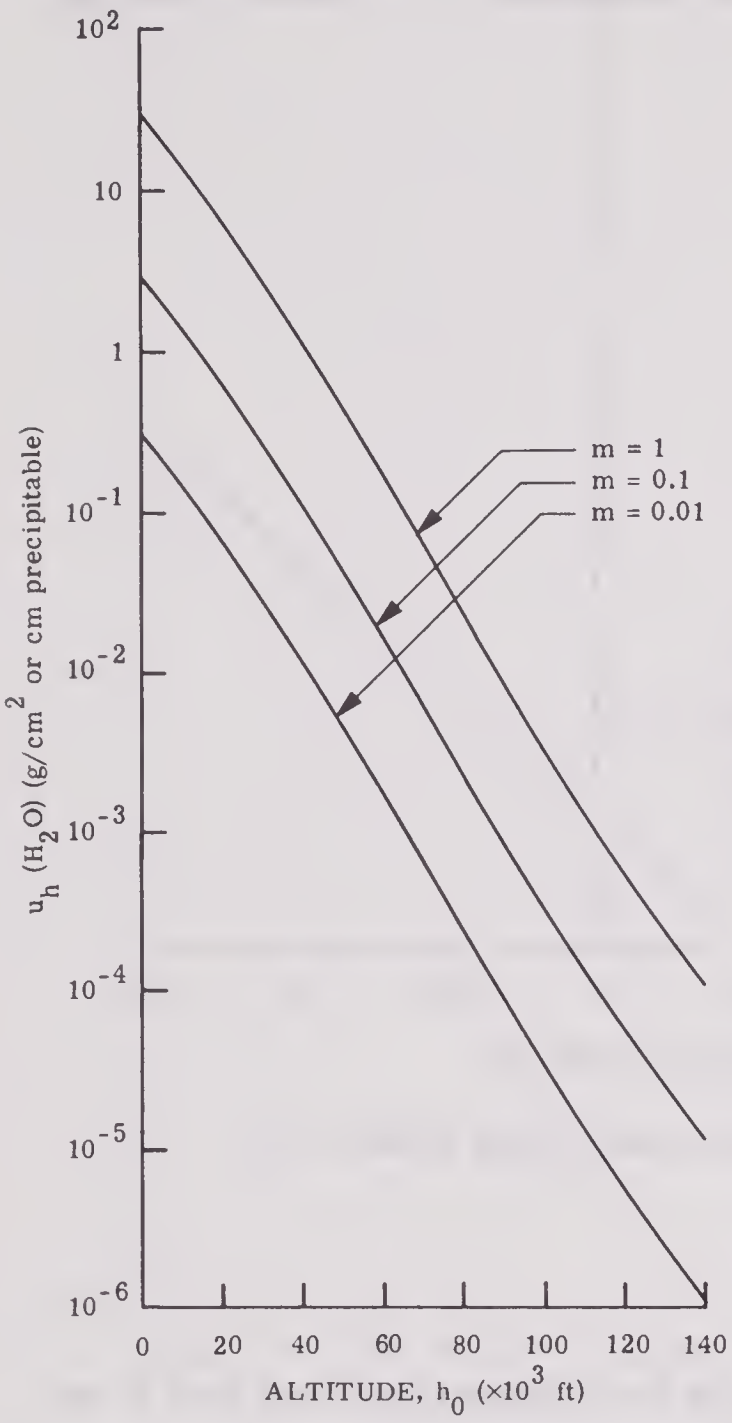


FIG. 6-135. Total horizontal equivalent optical path length for H_2O *vs.* altitude (for mixing ratios 1, 0.1, 0.01).

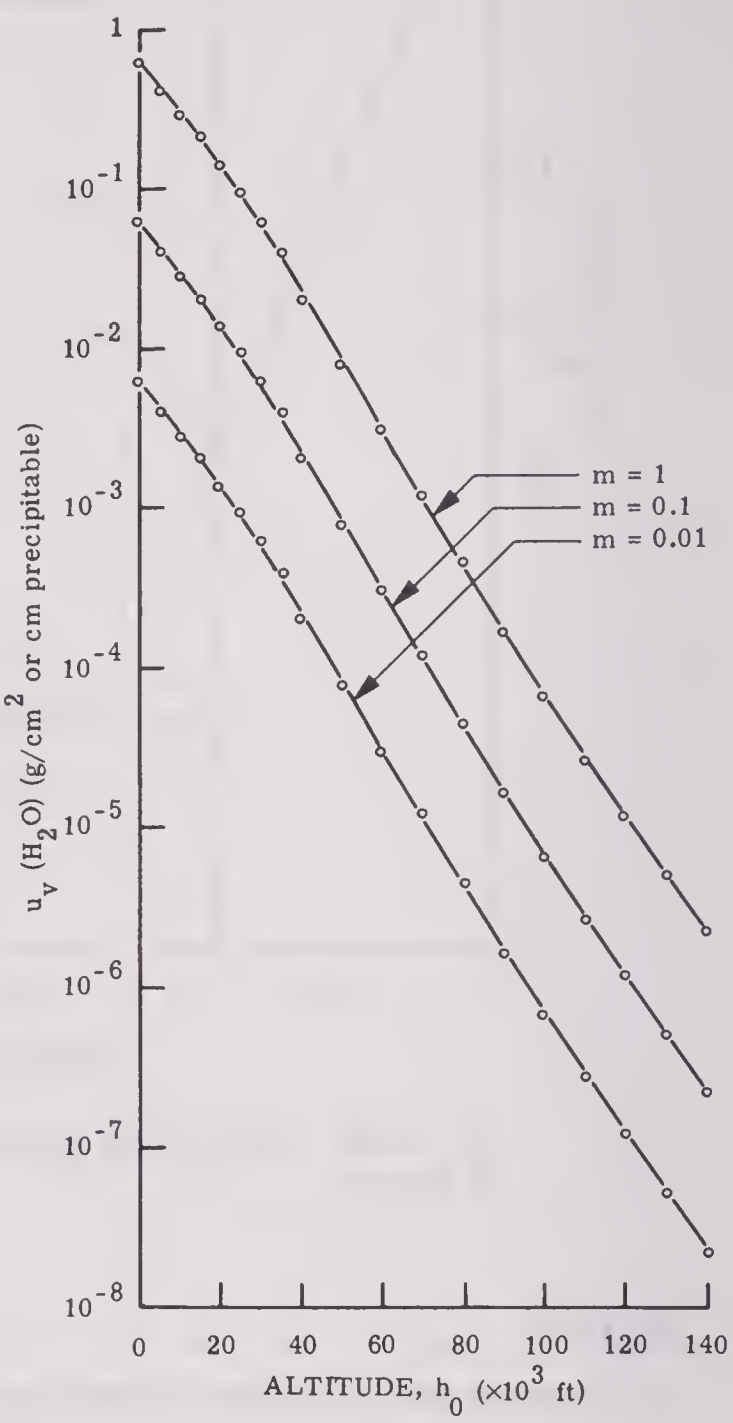


FIG. 6-136. Total vertical equivalent optical path length for H_2O *vs.* altitude (for mixing ratios 1, 0.1, 0.01).

STEPS 2-4 *for constant relative humidity.* When a constant relative humidity is assumed for the atmospheric path, the procedure is again similar to that given above with the exception that u_h should be determined from Fig. 6-137, u_v should be determined from Fig. 6-138, and H_0''' replaces H_0'' in Eq. (6-69), (6-70), and (6-71). Use Fig. 6-139 for the determination of H_0'''.

The reduced equivalent sea-level paths obtained using the above procedure are corrected for temperature and pressure as well as for the refraction of the atmospheric path. Values of the reduced sea-level path computed in this way will be quite accurate for temperate latitudes and may be used for the computation of slant-path transmissivities as outlined in the previous sections.

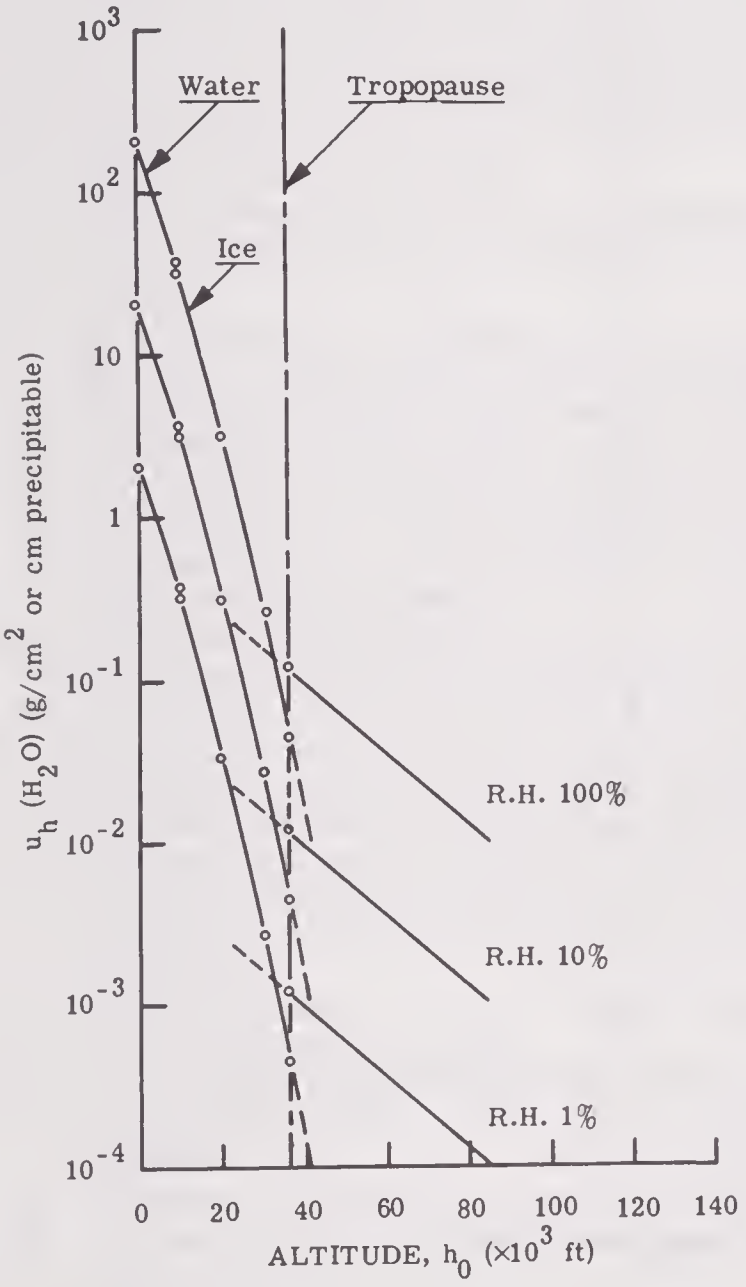


FIG. 6-137. Total horizontal equivalent optical path length for H_2O *vs.* altitude (for relative humidity 100%, 10%, 1%).

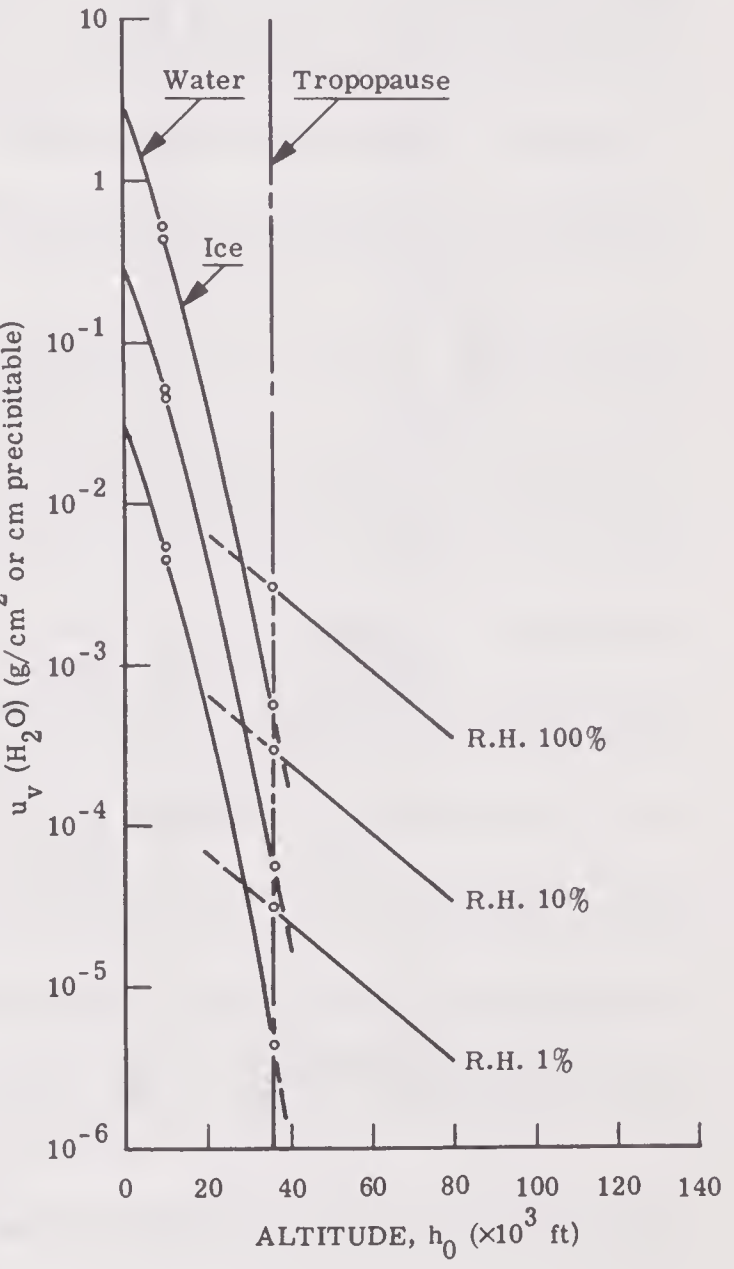


FIG. 6-138. Total vertical equivalent optical path length for H_2O *vs.* altitude (for relative humidity 100%, 10%, 1%).

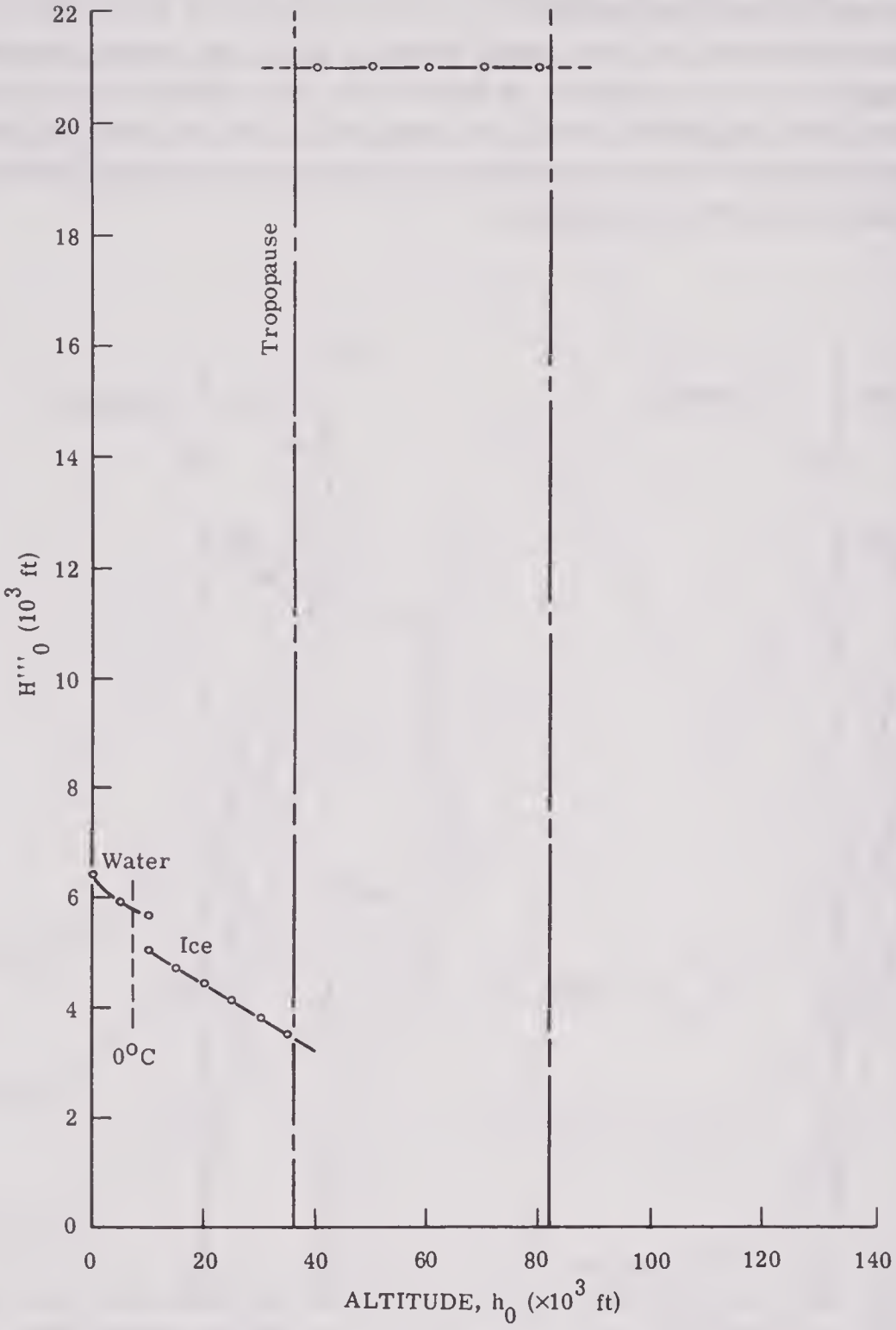


FIG. 6-139. Slant height for $P^2/T^{3/2}$.

References

1. A.E. Cole, A. Court, and A.J. Kantor, *Standard Atmosphere Revision to 90 Kilometers,* Report No. INAP-7, Rev. 2, Geophysics Research Directorate, AFCRL, Cambridge, Mass. (1961).
2. A. E. Cole, A. Court, and A. J. Kantor, *Supplemental Atmospheres,* Geophysics Research Directorate, AFCRL, Cambridge, Mass. (1963).
3. *Handbook of Geophysics-Revised Edition,* Macmillan Company, New York (1960).
4. C. E. Junge, *Atmospheric Chemistry, Advances in Geophysics, Vol. 4,* Academic Press, New York (1958).
5. M. Gutnick, "How Dry is the Sky?," *J. of Geo. Res.,* **66,** No. 9, 2867 (1961).
6. J. T. Houghton and J. S. Seeley, "Spectroscopic Observations of the Water Vapor Content of the Atmosphere," *Quart. J. Roy. Meteor. Soc.,* **86,** 358 (1960).
7. J. Yarnell and R. M. Goody, "Infrared Solar Spectroscopy in a High-altitude Aircraft," *J. Sci. Instr.,* **29,** 352 (1952).
8. A. W. Brewer, "Evidence for a World Circulation Provided by the Measurements of Helium and Water Vapor Distribution in the Stratosphere," *Quart. J. Roy. Meteor. Soc.,* **75,** 351 (1949).
9. G. M. B. Dobson, "Origin and Distribution of the Polyatomic Molecules in the Atmosphere," *Proc. Roy. Soc.,* London, **236,** No. 1205 187 (1956).
10. P. Goldsmith, "Some Aircraft and Surface Meteorology Observation Made at Khartoum," *Meteorol. Mag.,* **23,** 329 (1954).
11. N. C. Helliwell and J. K. Mackenzie, *Observations of Humidity, Temperature and Wind at Idris, 23rd May-2nd June 1956,* MRP 1024, Meteorological Research Committee, London (1957).
12. N. C. Helliwell, J. K. Mackenzie, and M. J. Kerley, *Further Observations of Humidity Up to 50,000 Feet, Made from an Aircraft of the Meteorological Research Flight in 1955,* MRP 976, Meteorological Research Committee, London (1956).
13. N. C. Helliwell, J. K. Mackenzie and M. J. Kerley, "Some Further Observations from Aircraft of Frost Point and Temperature up to 50,000 Feet," *Quart. J. Roy. Meteor. Soc.,* **83,** 257 (1957).
14. R.J. Murgatroyd, P. Goldsmith, and W.E.H. Hollings, "Some Recent Measurements of Humidity from Aircraft up to Heights of about 50,000 Feet over Southern England," *Quart. J. Roy. Meteor. Soc.,* **81,** 533 (1955).
15. D.M. Gates, D.G. Murcray, C.C. Shaw, and R.J. Herbold, "Near Infrared Solar Measurements by Balloons to Altitudes of 100,000 Feet," *J. Opt. Soc. Am.,* **48,** 1010 (1958).
16. D. G. Murcray, F. H. Murcray, W. J. Williams, and F. E. Leslie, *Water Vapor Distribution Above 90,000 Feet,* Scientific Report No. 5, Geophysics Research Directorate, Contract AF 19 (604)-2069, AFCRL, Cambridge, Mass. (1960).
17. N. Sissenwine and M. Gutnick, "Precipitable Water Along High Altitude Ray Paths," *Proc. of Infrared Information Symposium,* **5,** No. 2, 5 (1960).
18. H. J. Mastenbrook, and J. E. Dinger, *The Measurement of Water Vapor Distribution in the Stratosphere,* NRL Report 5551, U.S. Naval Research Laboratory, Wash., D.C., (1960) (ASTIA AD 247760).
19. M. J. Mastenbrook and J. E. Dinger, "Distribution of Water Vapor in the Stratosphere," *J. Geophysics,* **66,** 1437 (1961).
20. F. R. Barclay, M. J. Eliott, P. Goldsmith, and J. V. Jelley, "A Direct Measurement of the Humidity in the Stratosphere Using a Cooled-Vapor Trap," *Quart. J. Roy. Meteor. Soc.* **86,** 259 (1960).
21. F. Stauffer and J. Strong, *App. Optics,* **1,** (1962).
22. J. Strong, *Preliminary Report on Solar Observations from a U-2 Observatory,* The Johns Hopkins University, Baltimore, Md., (1960).
23. D. G. Murcray, F. H. Murcray, and W. J. Williams, *Distribution of Water Vapor in the Stratosphere as Determined from Infrared Absorption Measurements,* Scientific Report No. 1, University of Denver, Contract No. AF 19 (604)-7429, AFCRL Report No. 219, AFCRL, Cambridge, Mass., (1961).
24. R. K. McDonald, *Infrared Satellite Background Measurements, Part 1, Atmospheric Radiative Processes,* Final Report, Sept. 1961, The Boeing Company, Contract AF 19 (604)-7457 for Geophysics Research Directorate, AFCRL Report No. 1069, AFCRL, Cambridge, Mass.
25. T. L. Altshuler, *Infrared Transmission and Background Radiation by Clear Atmospheres,* Document No. 615D199, Dec. 1961, General Electric Company, Missile and Space Vehicle Department, Philadelphia, Pa.
26. M. Gutnick, *Mean Moisture Profiles to 31 Kilometers for Middle Latitudes,* Geophysics Research Directorate, AFCRL, Cambridge, Mass., (1962).
27. W. L. Godson, "Total Ozone and the Middle Stratosphere Over Arctic and Subarctic Areas in Winter and Spring," *Quart. J. Roy. Meteor. Soc.,* **86,** No. 369, 301 (1960).

28. E. O. Hurlburt, "Physics of the Upper Atmosphere," *Meteor. Res. Rev.* **3,** No. 17, 167 (1957).
29. R. M. Goody and W. T. Roach, "Determinations of the Vertical Distribution of Ozone from Emission Spectra," *Quart. J. Roy. Meteor. Soc.,* **82,** No. 352, 217 (1956).
30. H. Johansen, *On the Relation Between Meteorological Conditions and Total Amount of Ozone Over Tromso,* Polar Atmosphere Symposium, Part I, Pergamon Press, London, 187 (1958).
31. C. L. Mateer and W. L. Godson, "The Vertical Distribution of Atmospheric Ozone Over Canadian Stations from Umkehr Observations," *Quart. J. Roy. Meteor. Soc.,* **18,** No. 3, 512 (1960).
32. H. U. Dutsch, *Current Problems of the Photochemical Theory of Atmospheric Ozone,* paper given at Inter. Symp. on Chem. Reac. in the Lower and Upper Atmosphere, Stanford Research Institute, Stanford, California, April 18-20, 1961.
33. K. R. Ramanthan and R. N. Kulkarni, "Mean Meridional Distribution of Ozone in Different Seasons," Calculated from Umkehr Observations and Probable Vertical Transport Mechanisms, *Quart. J. Roy. Meteor. Soc.,* **86,** No. 368, 144 (1960).
34. H. K. Paetzold, *The Photochemistry of the Atmospheric Ozone-Layer,* paper given at Inter. Symp. on Chem. Reac. in the Lower and Upper Atmosphere, Stanford Research Institute, Stanford, California, April 18-20 (1960).
35. S. Verkateswaren, J. G. Moore, and A. J. Krueger, "Determination of the Vertical Distribution of Ozone by Satellite Photometry," *J. Geophys. Res.,* **66,** 1751, University of Calif., Berkeley, Calif., Naval Ordnance Test Station (1961).
36. J. N. Howard and J. S. Garing, *The Transmission of the Atmosphere in the Infrared,* GRD, AFCRL, Cambridge, Mass. (1962).
37. C. C. Junge, "Atmospheric Composition," *Handbook of Geophysics for Air Force Designers,* AFCRC, Cambridge, Mass. (1957).
38. A. J. Arnulf, J. Bricard, E. Cure and C. Veret, "Transmissions by Haze and Fog in the Spectral Region 0.35 to 10 Microns," *J. Opt. Soc. Am.,* **47,** 491 (1957).
39. H. J. Weickman and J. J. aufm Kampe, "Physical Properties of Cumulus Clouds," *J. Meterol.,* **10,** No. 204 (1953).
40. D. Deirmendjian, *Scattering and Polarization Properties of Polydispersed Suspension with Partial Absorption,* Rand Corp., Memorandum RM 3228 PR.
41. R. Penndorf, *The Vertical Distribution of Mie Particles in the Troposphere,* Geophysical Research Paper No. 25, AFCRC Tech. Report 54-5, AFCRC, Cambridge, Mass. (1954).
42. D. E. Burch, E.B. Singleton, and D. Williams, *App. Optics,* **1,** No. 359 (1962).
43. S. S. Penner, *Quantitative Molecular Spectroscopy and Gas Emissivities,* Addison-Wesley Publishing Co., Inc., Reading, Mass. (1959).
44. G. N. Plass and D. I. Fivel, *Astrophys. J.,* **117,** 225 (1953).
45. W. M. Elsasser, *Heat Transfer by Infrared Radiation in the Atmosphere, Harvard Meteor Studies No. 6,* Harvard University Press, Cambridge, Mass. (1942).
46. G. N. Plass, *J. Opt. Soc. Am.,* **48,** 690 (1958).
47. G. N. Plass, *J. Opt. Soc. Am.,* **50,** 868 (1960).
48. R. Ladenberg and F. Reiche, *Ann. Physik.,* **42,** 181 (1913).
49. G. N. Plass and D. I. Fivel, *Astrophys. J.,* **117,** 225 (1953).
50. W. M. Elsasser, *Phys. Rev.,* **54,** 126 (1938).
51. H. Mayer, *Methods of Opacity Calculations,* Los Alamos, LA-647 (1947).
52. R. M. Goody, *Quart. J. Roy. Meteor. Soc.,* **78,** 165 (1952).
53. L. D. Kaplan, *Proc. 1953 Toronto Meteor. Conf.,* 43 (1954).
54. P. J. Wyatt, V. R. Stull, and G. N. Plass, *J. Opt. Soc. Am.* (1962).
55. D. Q. Wark and M. Wolk, *Monthly Weather Review,* **88,** 249 (1960).
56. P. J. Wyatt, V. R. Stull, and G. N. Plass, *App. Optics,* **3** (1964); Aeronutronic Report U-1717, Aeronutronic Systems, Inc., Newport Beach, Calif. (1962).
57. V. R. Stull, P. J. Wyatt, and G. N. Plass, *App. Optics,* **3** (1964); Aeronutronic Report U-1718, Aeronutronic Systems, Inc., Newport Beach, Calif., 1962.
58. W. E. K. Middleton, *Vision Through the Atmosphere,* University of Toronto Press, Toronto, Canada (1952) Section 9.3.1.1.
59. W. E. K. Middleton, "The Effect of the Angular Aperture of a Telephotometer on the Telephotometry of Collimated and Non-Collimated Beams," *J. Opt. Soc. Am.,* **39,** No. 576 (1949).
60. H. S. Stewart and J. A. Curcio, "The Influence of Field-of-View on Measurements of Atmospheric Transmission," *J. Opt. Soc. Am.,* **42,** No. 801 (1952).
61. J. W. Tucker, *Computation of Singly-Scattered Radiation from a Distaat Source when the Angular Scattering Function is Known,* NRL Report 5260, U.S. Naval Research Laboratory, Wash., D.C. (1959).
62. K. Bulrich, "Die Streuung des Lichter in Truber Luft," *Optik,* **2,** No. 301 (1947).
63. R. Gans, "Wien-Harms Handbuch D. Exp.," *Physik,* **19,** No. 368.

64. J. A. Stratton and H. G. Houghton, "A Theoretical Investigation of the Transmission of Light Through Fog," *Phys. Rev.*, **38,** No. 159 (1931).
65. J. S. Stratton, *Electromagnetic Theory,* McGraw-Hill, New York (1941).
66. H. C. Van De Hulst, *Light Scattering by Small Particles,* Wiley, New York (1957).
67. H. G. Houghton and W. R. Cholker, "The Scattering Cross-Section of Water Crops in Air for Visible Light," *J. Opt. Soc. Am.,* **39,** No. 955 (1949).
68. *Tables of Scattering Functions for Spherical Particles,* NBS Applied Math., Series 4, National Bureau of Standards, Wash., D.C. (1949).
69. C-N Chu, G. C. Clark, and S. W. Churchill, *Tables of Angular Distribution Coefficients for Light Scattering by Spheres,* Engineering Research Institute, Department of Chemical and Metallurgical Engineering, The University of Michigan Press, Ann Arbor, Mich. (1957).
70. J. Bricard, "Etude de la constitution des nuages au sommet du Puy-de-Dome," *La Meterologie,* **15,** 83.
71. J. Bricard, "Lumiere diffusee en avante par une gotte d'eau spherique," *J. de Phys. et la Radium,* **4,** 57.
72. J. Bricard, "Reflexion, refraction et diffraction de la lumiere par une gotte d'eau spherique," *Ann. de Geophys.,* **2,** 231.
73. P. Kruse, L. McGlaughlin, and R. McQuistan, *Elements of Infrared Technology,* Wiley, New York (1962).
74. H. W. Yates and J. H. Taylor, *Infrared Transmission of the Atmosphere,* NRL Report 5453, U.S. Naval Research Laboratory, Wash., D.C. (1960).
75. M. G. Gibbons, "Transmission and Scattering Properties of a Nevada Desert Atmosphere," *J. Opt. Soc. Am.,* **51,** No. 633 (1961).
76. J. A. Curcio, G. L. Knestrick, and T. H. Cosden, *Atmospheric Scattering in the Visible and Infrared,* NRL Report 5567, U.S. Naval Research Laboratory, Wash., D.C. (1961) ASTIA AD 250945.
77. L. P. Granath and E. O. Hulburt, "The Absorption of Light by Fog," *Phys. Rev.,* **34,** No. 140 (1929).
78. J. A. Curcio and K. A. Durbin, *Atmospheric Transmission in the Visible Region,* NRL Report 5368, U.S. Naval Research Laboratory, Wash., D.C.
79. F. Löhle, "Über die Lichtzerstreuung im Nebel," *Phys. Zeits.,* **45,** No. 199 (1944).
80. M. Wolff, "Die Lichttechnischen Eigenschaften des Nebels," *Das Licht,* **8,** No. 105 and 128 (1938).
81. *Measurement of Atmospheric Attenuation aboard USAS AMERICAN MARINER,* Project DAMP Progress Report, Barnes Eng. Co., Stamford, Conn.
82. H. B. Glenn, *Light Transmission Through an Apparently Clear Atmosphere,* Raytheon Co., Santa Barbara Research Operations, Santa Barbara, Calif. (1960).
83. W. M. Protheroe, *Preliminary Report on Stellar Scintillation,* Scientific Report No. 4, Ohio State University, Contract No. AF 19 (604)-41, AFCRL, Cambridge, Mass., (1954) ASTIA AD 56040.
84. G. Keller, W. M. Protheroe, P. E. Barnhard and J. Galli, *Investigation of Stellar Scintillation and the Behavior of Telescopic Images,* Final Report, Contract No. AF 19 (604)-1409, AFCRC, Report No. TR-57-186, AFCRC, Cambridge, Mass. (1956) ASTIA AD 117279.
85. E. Goldstein, *The Measurement of Fluctuating Radiation Components in the Sky and Atmosphere, Part 1,* NRL Report N-3462, U.S. Naval Research Laboratory, Wash., D.C. (1949) ASTIA ATI-7149.
86. E. Goldstein, *The Measurement of Fluctuating Radiation Components in the Sky and Atmosphere, Part 2,* NRL Report 3710, U.S. Naval Research Laboratory, Wash., D.C. (1950) ASTIA PB-102617.
87. F. R. Bellaire and F. C. Elder, *Scintillation and Visual Resolution Over the Ground,* Report of Project MICHIGAN, 2900-134-T, The University of Michigan, Willow Run Lab., Ann Arbor, Mich. (1960) ASTIA AD 245118.
88. B. N. Bullock, G. M. Smith, and R. P. Borofka, *Atmospheric Boil Measurements,* ITT Federal Laboratories, San Fernando, Calif.
89. F. Benford, "Duration of Intensity of Sunshine—Part I - General Considerations and Corrections," *Illuminating Engineering,* **42,** 527, General Electric Co. (1947).
90. M. Migeotte, L. Nevin, and J. Swensson, *The Solar Spectrum from 2.8 to 23.7 Microns, Part I. Photometric Atlas,* University of Liege, Contract AF 61 (514)-432, Phase A, Part I, Geophysics Research Directorate, AFCRC, Cambridge, Mass. ASTIA AD 210043.
91. M. Migeotte, L. Nevin, and J. Swensson, *The Solar Spectrum from 2.8 to 23.7 Microns, Part II. Measures and Identifications,* University of Liege, Contract AF 61(514)-432, Phase A, Part II, Geophysics Research Directorate, AFCRC, Cambridge, Mass. ASTIA AD 210044.

92. M. Migeotte, L. Nevin, and J. Swensson, *An Atlas of Nitrous Oxide, Methane and Ozone Infrared Absorption Bands, Part I. The Photometric Records,* University of Liege, Contract AF 61(614)-432, Phase B, Part I, Geophysics Research Directorate, AFCRC, Cambridge, Mass. ASTIA AD 210045.
93. M. Migeotte, L. Nevin, and J. Swensson, *An Atlas of Nitrous Oxide, Methane and Ozone Infrared Absorption Bands, Part II. Measures and Identifications,* University of Liege, Contract AF 61(514)-432, Phase B, Part II, Geophysics Research Directorate, AFCRC, Cambridge, Mass. ASTIA AD 210046.
94. J. N. Howard and J. S. Garing, *Infrared Atmospheric Transmission: Some Source Papers on the Solar Spectrum from 3 to 15 Microns,* Air Force Surveys in Geophysics No. 142, AFCRL Report No. 1098, Dec. 1961, Geophysics Research Directorate, AFCRL, Cambridge, Mass.
95. J. N. Howard, "Atmospheric Transmission in the 3 to 5 Micron Region," *Proc. IRIS,* **2,** 59-75 (1957).
96. J. N. Howard, "Atmospheric Transmission in the 8 to 13 Micron Region," *Proc. of the Symposium on Optical Radiation from Military Airborne Targets,* Final Report No. AFCRL-TR-58-146, AFCRL, Cambridge, Mass., Contract No. AF 19(604)-2451, Haller, Raymond and Brown, Inc., State College, Pa. ASTIA AD 152411; *see also* [7].
97. O. C. Mohler, A. K. Pierce, P. R. McMath, and L. Goldberg, *Atlas of the Solar Spectrum from 0.84 to 2.52 Microns,* University of Michigan Press, Ann Arbor, Mich. (1950).
98. O. C. Mohler, *Table of Solar Spectrum Wavelengths from 1.20 to 2.55 Microns,* University of Michigan Press, Ann Arbor, Michigan (1955).
99. J. H. Shaw, R. M. Chapman, J. N. Howard, and M. L. Oxholm, "A Grating Map of the Solar Spectrum from 3.0 to 5.0 Microns," *Astrophys. J.,* **113,** No. 2, (1951) *see also* [7].
100. J. H. Shaw, M. L. Oxholm, and H. H. Classen, "The Solar Spectrum from 7 to 13 Microns," *Astrophys. J.,* **116,** No. 3 (1952) *see also* [7].
101. W. W. Talbert, H. A. Templin, and R. E. Morrison, *Quantitative Solar Spectral Measurements at Mt. Chacaltaya (17,100 ft.),* U.S. Naval Ordnance Laboratory, White Oak, Maryland; See also *J. Opt. Soc. Am.,* **47,** 156 (1957).
102. J. E. Seeley, J. T. Houghton, T. S. Moss, and N. D. Hughes, "Solar Spectrum from 1 to 6.5 Microns at Altitudes Up to 15 KM," *Phil. Trans. Roy. Society.*
103. C. B. Farmer and S. J. Todd, *Reduced Solar Spectra 3.5 to 5.5 Microns,* Report DP 927 E.M.I. Electronics, Hayes, Middlesex, England (1961).
104. D. E. Burch and D. Williams, *Infrared Absorption by Minor Atmospheric Constituents,* The Ohio State University Research Foundation, Scientific Report No. 1, Contract No, AF 19(604)-2633, Geophysics Research Directorate, AFCRL Report No. TN-60-674, AFCRL, Cambridge, Mass. (1960) ASTIA AD 246921.
105. D. E. Burch, D. Gryvnak, and D. Williams, *Infrared Absorption by Carbon Dioxide,* The Ohio State University Research Foundation, Scientific Report No. 11, Contract No. AF 19(604)-2632, Geophysics Research Directorate, AFCRL Report No. 255, AFCRL, Cambridge, Mass. (1960) ASTIA AD 253435.
106. D. E. Burch, E. B. Singleton, W. L. France, and D. Williams, *Infrared Absorption by Minor Atmospheric Constituents,* The Ohio State University Research Foundation, Final Report, Contract No. AF 19(604)-2633. Geophysics Research Directorate, AFCRL Report No. 412, AFCRL, Cambridge, Mass. (1960) ASTIA AD 256952.
107. J. N. Howard, D. E. Burch, and D. Williams, *Near-Infrared Transmission Through Synthetic Atmospheres,* AFCRL Report No. TR-55-213, Geophysics Research Directorate, AFCRL, Cambridge, Mass. (1955).
108. H. W. Yates and J. H. Taylor, *Infrared Transmission of the Atmosphere,* NRL Report 5453, U.S. Naval Research Laboratory, Wash., D.C. (1960) ASTIA AD 240188.
109. T. Elder and J. Strong, "The Infrared Transmission of the Atmospheric Windows," *Journal of the Franklin Institute,* **255,** No. 3, 189 Phila., Pa. (1953).
110. C. B. Farmer, *The Transmission of Infrared Radiation Through Cloud,* Report No. DP. 841, E.M.I. Electronics, Ltd., Hayes, Middlesex, England, (1960).
111. T. L. Altshuler, *Infrared Transmission and Background Radiation by Clear Atmospheres,* General Electric Company, Missile and Space Vehicle Department, Valley Forge, Pa., Docu-61SD199 (1961).
112. G. N. Plass, and D. I. Fivel, *J. Meteorol.,* **12,** 191 (1955).
113. G. N. Plass, *App. Optics* **2,** 515 (1963).
114. G. N. Plass, *J. Opt. Soc. Am.,* **42,** 677 (1952).
115. R. O'B. Carpenter, J. A. Wight, A. Quesda, and R. E. Swing, *Predicting Infrared Molecular Attenuation for Long Slant Paths in the Upper Atmosphere,* AFCRC Report No. TN-58-253, AFCRC, Cambridge, Mass. (1957).
116. A. R. Curtis, *Quart. J. Roy. Meteor. Soc.,* **78,** 638 (1952).

117. W. L. Godson, *J. Meteorol.*, **12**, 272 (1955).
118. A. Thomson and M. Downing, *Earth Background and Atmospheric Transmission in the 2 to 5 Micron Region,* Convair Report Ph-069-M (1960).
119. A. S. Zachor, *Near Infrared Transmission Over Atmospheric Slant Paths,* Report R-328, **2**, Massachusetts Institute of Technology, Cambridge, Mass., Contract AF 33(616)-6046 (1961).
120. A. E. S. Green and M. Griggs, *Appl. Optics* **2**, 561 (1963).
121. G. A. Morton and G. M. Weyl, *Water Vapor and Carbon Dioxide Absorption in the Spectral Region Around 2.7 Microns,* Special Report No. 2075, Aerojet-General Corp., Azusa, Calif. (1961).
122. L. R. Megill and P. M. Jamnick, *J. Opt. Soc. Am.*, **51**, No. 1294 (1961).
123. G. N. Plass, *Transmittance of Carbon Dioxide and Water Vapor over Stratospheric Slant Paths,* Aeronutronic Report, Aeronutronic Systems, Inc., Newport Beach Calif. (1962); *Appl. Optics,* **3** (1964).

Chapter 7

OPTICAL COMPONENTS

William L. Wolfe

The University of Michigan

CONTENTS

7.1. Lenses 282
7.2. Mirrors 285
7.3. Catadioptric Systems 285
7.4. Interference Filters 286
7.4.1. Introduction 286
7.4.2. Terminology 286
7.4.3. General Theory of Interference Filters 288
7.4.4. Long-Wave Pass Interference Filters 290
7.4.5. Short-Wave Pass Interference Filters 290
7.4.6. Bandpass Interference Filters 290
7.4.7. "Square-Band" Interference Filters 291
7.4.8. The Filter Matrix 291
7.4.9. The Herpin Equivalent Layer 293
7.4.10. Analogies with Transmission-Line Theory 293
7.4.11. Effects of Angle of Incidence 293
7.4.12. Effects of Temperature 294
7.4.13. Substrates and Films 294
7.5. Christiansen Filters 295
7.6. Selective Reflection Filters 297
7.7. Selective Refraction Filters 298
7.8. Polarization Interference Filters 298
7.9. Commercially Available Filters 299
7.10. Absorption Filters 306
7.11. Prisms 307
7.11.1. Dispersing Prisms 308
7.11.2. Deviating Prisms 308
7.11.3. Prism Materials 309
7.12. Diffraction Gratings 309
7.12.1. Blazed Gratings 310
7.12.2. Concave Gratings 311
7.12.3. Ebert-Fastie Plane Grating Mountings 311
7.12.4. Concave Grating Mountings 311
7.12.5. Production of Gratings 313

7. Optical Components

7.1. Lenses

Infrared lenses and lens systems are usually custom made and are often designed by manufacturers. Stock items which are available are made from the more rugged materials such as Irtran, arsenic trisulfide, silicon, germanium, and optical glass; these materials are described in Chapter 8. Optical glass lenses are not discussed because they are easily obtained in almost any aperture and focal length.*

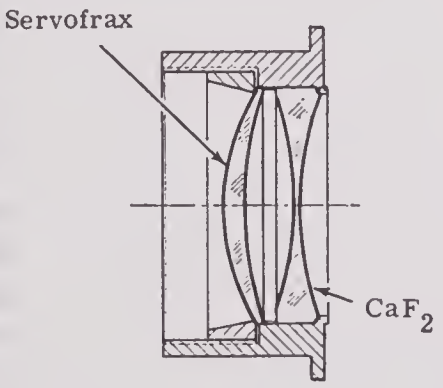
Tables 7-1 through 7-3 and Figs. 7-1 through 7-5 show various types of lenses and lens systems supplied by manufacturers for infrared applications.

The theory of refracting optics including thin and thick lenses and multiple lens systems is covered in detail in Chapter 9.

TABLE 7-1. SINGLE-ELEMENT LENSES SUPPLIED BY EASTMAN KODAK COMPANY [1].
(All are Irtran-2, Meniscus, measurements are in inches)

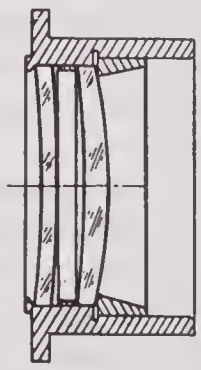
Range Δλ	*Nominal Focal Length* *f*	*Equivalent Focal Length* *e.f.l.*	*Back Focal Length* *b.f.l.*	*Focal Ratio* *f*/no	*Circle of Confusion* *d*†	*Distance C. of C. from Vertex*	*Lens No.*
1.5-10	1	0.99*	0.90*	1	0.028	0.84	IR-100
		1.04	0.92		0.028	0.84	
1.5-10	2	1.99	1.80	1	0.060	1.66	IR-200
		2.12	1.91		0.055	1.77	
3-10	1	1.00	0.86	1	0	0.86	IR-101
		1.05	0.90		0.011	0.87	
4.26-10	2	1.92	1.68	1	0.001	1.68	IR-201
		2.00	1.75		0.001	1.75	
4.26-10	3	3.00	2.72	1	0.001	2.72	IR-301
		3.13	2.84		0.001	2.84	

*The first number is for the short-wavelength end of the range; the second is for the long wavelength end.
†To the 5% intensity point; measurements made 2.0 to 4.5 μ.



(a) (b)

FIG. 7-1. Servocon ® achromats by SCA [2].



All Lenses Germanium

FIG. 7-2. SCA doublet [2].

*For extensive lists see the catalogs of appropriate companies listed in the Optical Industry Directory.

TABLE 7-2. SINGLE-ELEMENT INFRARED LENSES SUPPLIED BY SERVO CORPORATION OF AMERICA [2].

Lens Shape	*Range* $\Delta\lambda$ (μ)	*Design* λ λ_d (μ)	*Nominal Focal Length f* (in.)	*Effective Focal Length at* λ_d (in.) *e.f.l.*	*Back Focal Length at* λ_d (in.) *b.f.l.*	*f*/no (in.)	*Circle of Confusion dia. at* λ_d (in.)	*Circle of Confusion over* $\Delta\lambda$ (in.)	*Material*
Convex-Convex	0.7-1.5	1.1	4.0	4.0	3.683	4	0.006	0.011	Fused quartz
			2.0	2.0	1.753	2	0.024	0.023	Fused quartz
			4.0	4.0	3.523	2	0.047	0.055	Fused quartz
	0.7-2.0	1.4	2.0	2.0	1.748	2	0.027	0.030	CaF_2
			4.0	4.0	3.491	2	0.053	0.060	CaF_2
			4.0	4.0	3.665	4	0.007	0.011	CaF_2
	0.7-3.0	1.8	2.0	2.0	1.759	2	0.022	0.028	BaF_2
			4.0	4.0	3.713	4	0.006	0.011	BaF_2
			4.0	4.0	3.536	2	0.044	0.055	BaF_2
Plano-Convex	2.0-5.0	5.0	3.0	–	–	0.8	–	–	Servofrax
		5.0	2.0	–	–	2.0	–	–	Servofrax
		5.0	0.75	–	–	0.8	–	–	Servofrax
		4.0	0.57	–	–	0.65	–	–	Servofrax
Equi-convex	2.0-5.0	4.0	0.55	–	–	0.63	–	–	Servofrax
		6.0	2.0	–	–	1.0	–	–	Servofrax
		6.0	4.0	–	–	1.3	–	–	Servofrax
Meniscus	1.0-2.0	1.5	2.0	2.0	1.853	2	0.018	0.022	MgO
			4.0	4.0	3.877	4	0.005	0.003	MgO
	2.0-5.0	3.5	0.75	0.75	0.647	1	0.024	0.027	Servofrax
			1.5	1.5	1.378	3	0.003	0.005	Servofrax
			2.0	2.0	1.905	4	0.0006	0.0025	Servofrax
			2.0	2.0	1.905	2	0.005	0.009	Servofrax
			3.6	3.59	3.464	1.2	0.044	0.051	Servofrax
			4.0	4.0	3.835	4	0.001	0.005	Servofrax
			4.0	4.0	3.835	2	0.010	0.017	Servofrax
			5.5	5.5	5.094	1	0.176	0.190	Servofrax
			8.0	8.0	7.671	4	0.0025	0.010	Servofrax
			8.0	8.0	7.671	2	0.021	0.037	Servofrax
			9.0	9.0	8.796	3	0.019	0.030	Servofrax
			14.3	14.29	13.889	1.9	0.041	0.066	Servofrax
			35.9	35.85	35.594	4.5	0.010	0.041	Servofrax
			35.9	35.85	35.594	3.6	0.020	0.056	Servofrax
	2.0-11.0	6.5	2.0	2.0	1.926	2	0.003	0.006	Silcon
			4.0	4.0	3.876	2	0.006	0.012	Silcon
			–	4.0	3.899	4	0.001	0.005	Silcon
	6.0-10.0	8.0	2.0	2.0	1.916	2	0.005	0.009	Servofrax
			4.0	4.0	3.856	2	0.010	0.019	Servofrax
			4.0	4.0	3.886	4	0.001	0.006	Servofrax
	6.0-16.0	11.0	2.0	2.0	1.926	2	0.0025	0.003	Germanium
			4.0	4.0	3.880	2	0.005	0.006	Germanium
			4.0	4.0	3.901	4	0.0005	0.001	Germanium

TABLE 7-3. MULTI-ELEMENT LENS SYSTEMS SUPPLIED BY SERVO CORPORATION OF AMERICA [2].

Range $\Delta\lambda$ (μ)	*Design Wavelength* λ_d (μ)	*Nominal Focal Length* f (in.)	*Back Focal Length b.f.l.* (in.)	*Aperture* (f/no)	*Circle of Confusion* λ_d (in.)	*Radius of Petzval* (in.)	*Field of View* (degrees)	*Angular Resolution* (mrad)
2.0-5.0 (Fig. 7-1a)	3.5	2.0	–	1.2	0.004	5.60	0 1.0 2.0 4.0 6.0	2.00 2.24 2.63 3.87 5.72
2.0-5.0 (Fig. 7-1a)	3.5	4.0	–	1.2	0.008	11.20	0 1.0 2.0 4.0 6.0	2.0 2.24 2.63 3.87 5.72
2.0-5.0 (Fig. 7-1a)	3.5	8.0	–	2.0	0.003	23.5	0 1.0 2.0 4.0 6.0	0.38 0.43 0.50 0.75 0.90
3.5-5.0 (Fig. 7-1b)	4.5	4.75	3.56	1.2	0.0026	23.5	0 1.0 2.0 4.0 6.0	0.55 0.69 0.90 1.56 2.52
6.0-16.0 (Fig. 7-2)	11.0	4.0	–	3	–	–	0 9.0	Diffraction limit 0.6
1.2-4.0 (Fig. 7-3)	2.5	3.0	–	2	–	–	0 6.0	0.40 0.80
1.2-2.0 (Fig. 7-4)	1.6	3.6	–	1.6	–	–	0 6.0 9.0	1.1 1.6 2.2
6.0-16.0 (Fig. 7-5)	11.0	2.2	–	1.5	–	–	0	Diffraction limit

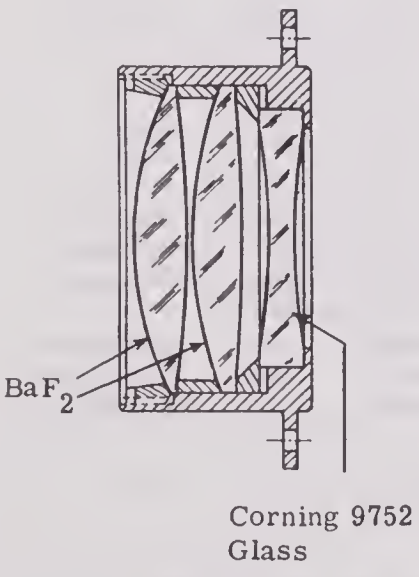


FIG. 7-3. SCA three-element achromat [2].

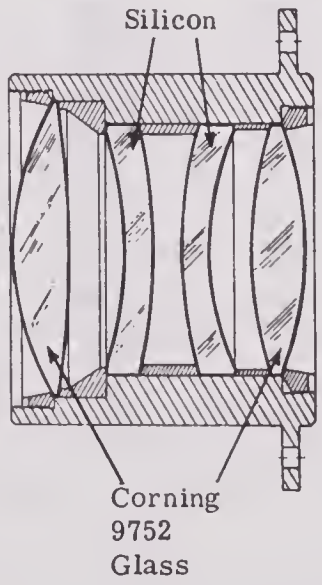


FIG. 7-4. SCA four-element achromat [2].

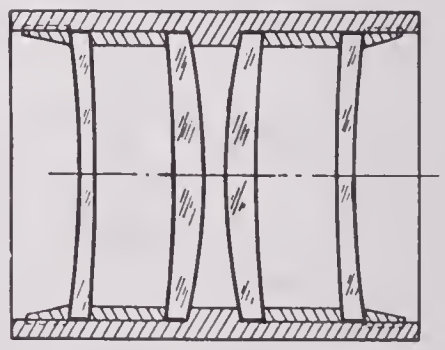


FIG. 7-5. SCA relay lens [2].

7.2. Mirrors

Mirrors in infrared systems, used as reflectors and as image-forming devices, often provide the additional function of filtering. Most infrared-system mirrors are front surfaced; however, they are back surfaced when used additionally as correctors. Mirrors for infrared instrumentation are generally not commercially available, but are custom designed for a particular application. Chapter 10 presents the theory and equations of reflecting optics. Chapter 8 discusses optical surface coatings, including reflective and protective coatings, filter mirrors, and replica mirrors.

7.3. Catadioptric Systems

Combined reflecting-refracting, or catadioptric, systems provide correction of aberrations over relatively wide angular fields. Table 7-4 and Figs. 7-6 through 7-10 show various commercially available catadioptric systems.

Theory and equations of spherical and aspherical concentric optical systems are given in detail in Chapter 10.

TABLE 7-4. CHARACTERISTICS OF CATADIOPTRIC SYSTEMS SUPPLIED BY SERVO CORPORATION OF AMERICA (S) AND BARNES ENGINEERING COMPANY (B) [2,3].

Optical System	*Wavelength Range* (μ)	*Design Wavelength* (μ)	*Nominal Focal Length* (in.)	*Aperture* (*f*/no)	*Primary Diameter* (in.)	*Lens Type*	*Field of View* (degrees)	*Angular Resolution* (mrad)	*Supplier*
Maksutov Folded (Fig. 7-6)	0.4-2.0	1.2	10.0	1.6	–	Maksutov quartz corrector	0 6	0.88 2.61	S
Bouwers-Maksutov (Fig. 7-7)	1.0-4.0	2.5	9.6	1.6	–	Maksutov As_2S_3 corrector	0 4	1.5 1.5	S
Catadioptric (Fig. 7-8)	0.3-0.6	0.45	10.1	3.3	–	Maksutov quartz corrector	0 6	0.15 0.15	S
Folded (Fig. 7-9)	1.0-3.0	–	6.0	1.8	3.3	Corning 9752[a]	2	1 at 0° 8 at 10°	B
Folded (Fig. 7-9)	1.0-3.0	–	4.1	1.4	5.3	Fused silica	20	1 at 0° 8 at 10°	B

[a]Germanate glass.

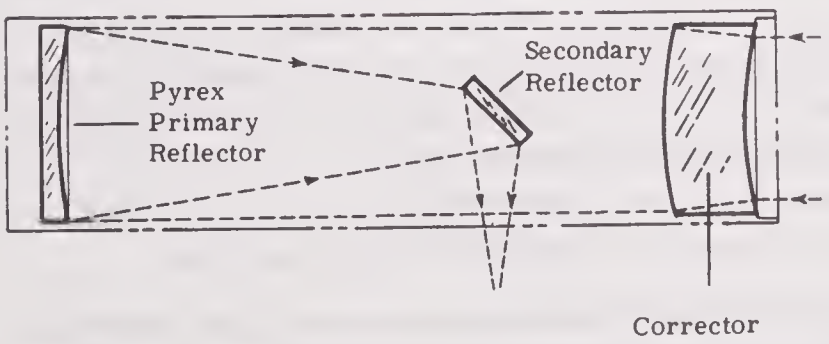


FIG. 7-6. Maksutov folded optical system supplied by SCA [2].

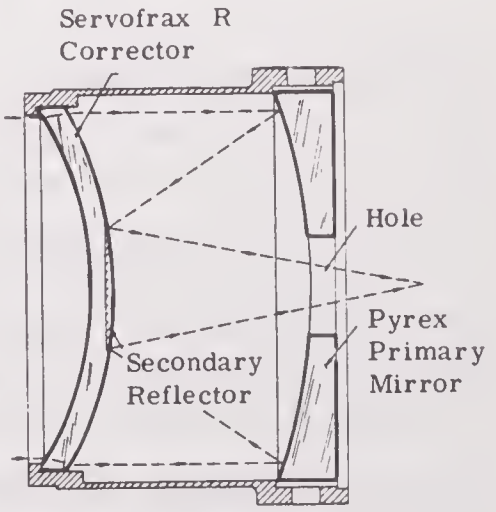


FIG. 7-7. Bouwers-Maksutov optical system supplied by SCA [2].

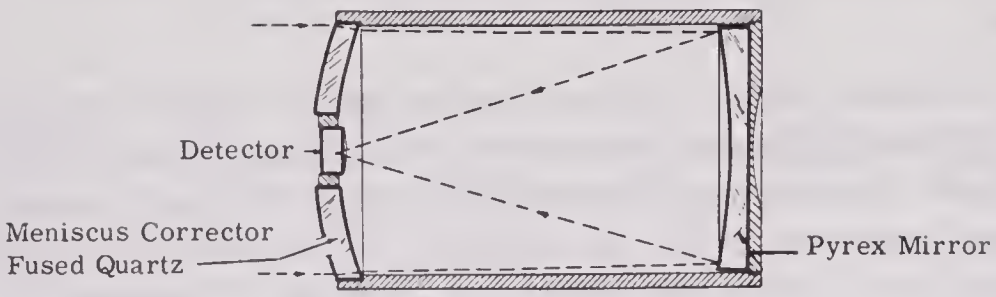


FIG. 7-8. Catadioptric optical system supplied by SCA [2].

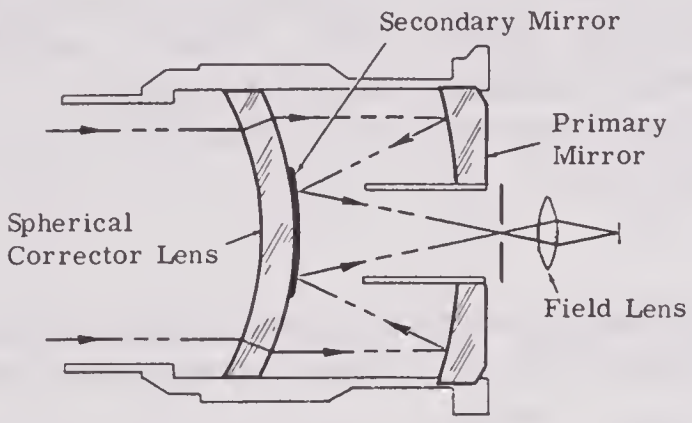


FIG. 7-9. Barnes high-speed objective system [3].

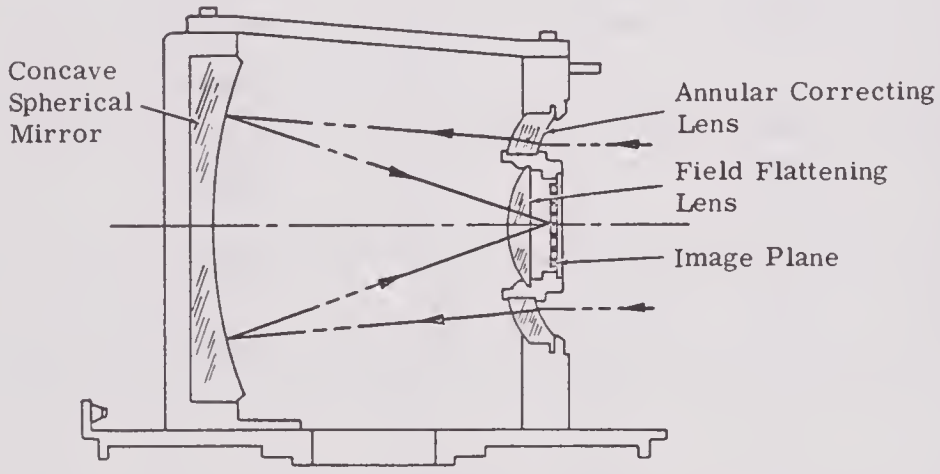


FIG. 7-10. Barnes wide-field objective system [3].

7.4. Interference Filters

7.4.1. Introduction. Filters are classified by either their transmission characteristics or the physical phenomena upon which their action is based. A long-wave pass filter transmits all radiation with wavelengths greater than the specified value; a short-wave pass filter passes all radiation with wavelengths shorter than the specified value; a bandpass filter transmits only between two wavelengths.

Some of the physical phenomena that determine filter action are selective reflection and refraction, scattering, polarization, interference, and selective absorption.

7.4.2. Terminology. The description of filters—and even curves of their transmission—has not been standardized. The terms given below have received some measure of acceptance, but for precise knowledge of characteristics nothing can substitute for a transmittance curve. For the following definitions, refer to Fig. 7-11.

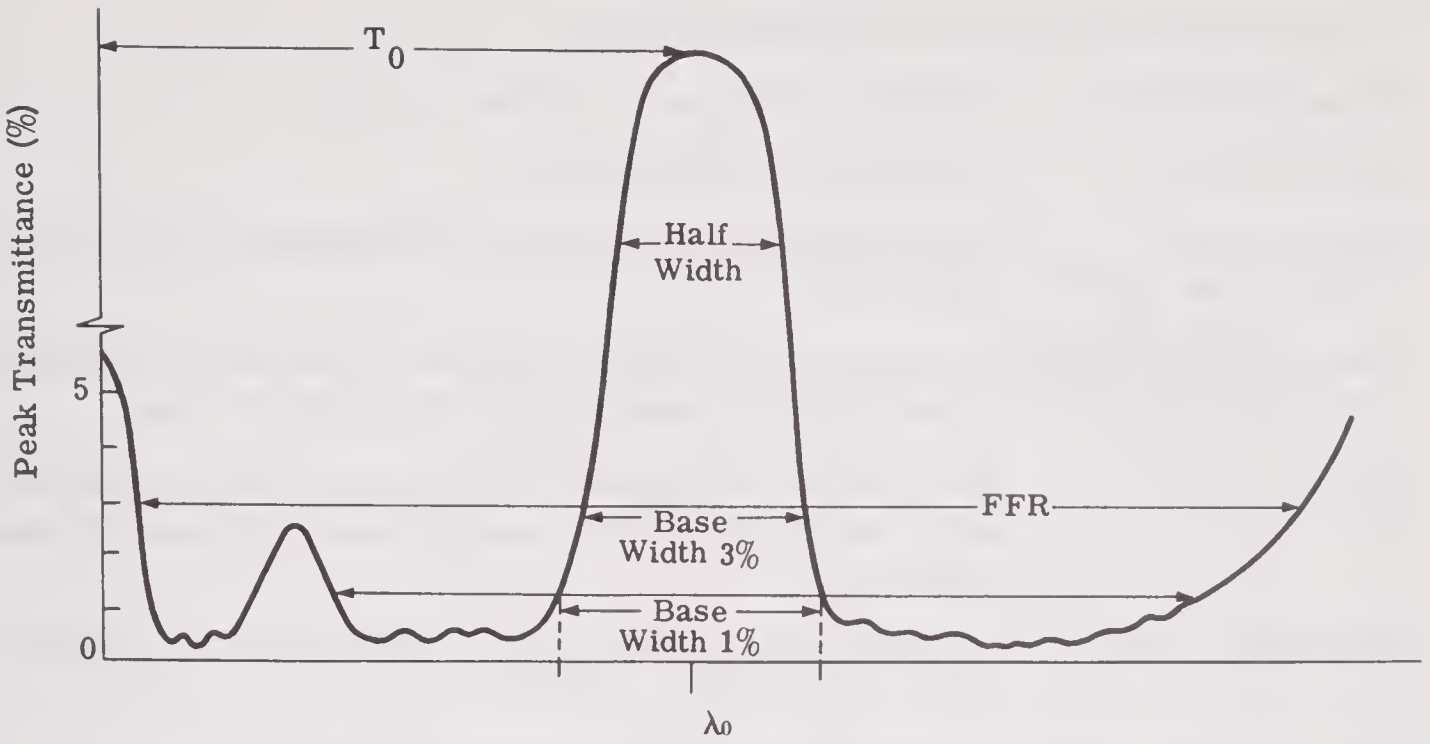


FIG. 7-11. Illustrating filter nomenclature.

Passband: The primary wavelength interval of transmission of a pass filter.

Stop Band: The primary region of reflection, absorption, etc., of a rejection filter.

Background Region: The region of low transmission of a pass filter.

Center Wavelength (λ_0): The wavelength at the center of the passband. For interference filters, λ_0 is usually the mean of the long- and short-wave cutoffs.

Halfwidth (HW): The width of the passband at 50% peak transmission—often expressed as a percentage of λ_0.

Basewidth (BW): The width of the passband of 3% or 1% (or some other small percentage) of peak transmittance. It is more precise and useful to specify the 3% width, etc., as $(\Delta\lambda)_{0.03}$; thus HW becomes $(\Delta\lambda)_{0.5}$.

Peak Transmittance (T_0): The maximum transmittance in the passband. For interference filters this is often specified as a percent of the uncoated substrate.

Substrate Transmittance (T_s): Substrate transmittance.

Free Filter Range (FFR): The wavelength interval over which the background is less than a specified amount except where the passband is.

Cutoff and Cuton Wavelengths (λ_c): The wavelengths of the limits of the passband.

Slope: The linear approximation to the cuton or cutoff slope, expressed as the ratio $(\lambda_{0.8} - \lambda_c)/\lambda_c$, where $\lambda_{0.8}$ is the wavelength of 80% transmittance.

The user of filters should take care to understand the nomenclature used by the individual manufacturers. In particular he should note whether T_0 is given in terms of the ratio of the filter transmission to the substrate transmission, whether λ_c is given as a 5%, 3%, or 1% cutoff, and whether FFR is specified for 1%, 0.1%, etc.

7.4.3. General Theory of Interference Filters.

Phase Difference (δ): The phase difference in radians is $2\pi\sigma nt$.

Wavenumber (σ): The reciprocal of wavelength.

Dimensionless Wavenumber (g): Some design (set) wavelength divided by the variable wavelength, λ_0/λ.

Quarter-Wave Optical Thickness (QWOT): The thickness of a layer measured in the number of quarter waves of a design wavelength. Thus, $QWOT = 4nt$.

H *and* L *Layers*: Quarter-wave thicknesses of the higher index in a stack of layers are often written H; quarter-wave thickness of lower index as L.

Quarter-Wave Stack: A periodic array of alternating H and L layers, *viz.*, $HLHLHL = (HL)^3$ or $LHLHLH = (LH)^3$.

High-Reflectance Zone: Region of high reflectance as shown in Fig. 7-12.

Quarter-wave stacks are the most simple from the design standpoint. On a frequency scale they have the symmetric transmission structure shown in Fig. 7-12. The width of the high reflectance zone is given by

$$\Delta g = \frac{4}{\pi} \arcsin\left(\frac{n_H - n_L}{n_H + n_L}\right)$$

The maximum reflectivity is given by

$$R_{max} = \frac{P + P^{-1} - 2}{P + P^{-1} + 2}$$

$$P = \left[\frac{n_l}{n_{l-1}} \frac{n_{l-2}}{n_{l-3}} \cdots \frac{n_2}{n_1}\right]^2 \frac{n_0}{n_s} \qquad l \text{ even}$$

$$P = \left[\frac{n_l}{n_{l-1}} \cdots n_1\right]^2 \frac{1}{n_0 n_s} \qquad l \text{ odd}$$

For $P >> P^{-1}$,

$$R_{max} = \frac{P - 4}{P}$$

The theoretical curve given in Fig. 7-12 is useful for beginning calculations. Figure 7-13 shows the transmission of a quarter-wave stack over a larger range of g.

Stacks with unequal optical thickness ratios can also be very useful. The 2:1 stack has the configuration

$$LLH\ LLH\ LLH\ \cdots = (LLH)^m$$

The first-order high-reflectance zone occurs at $g = 1$, when LLH is $\lambda_0/2$. The second-order zone occurs when $g = 2$ and LLH is λ_0. There is no high reflectance when $g = 3$ because LLH is $3\lambda_0/2$; thus H is a half wave and LL is a full wave, and all LLH's are absentee layers.

The general p:q stack can be analyzed in a similar way. The following features may be useful:

1. The high-reflectance zone may not be a center of symmetry
2. The high-reflectance zone of a quarter-wave stack is wider than other p:q stacks.

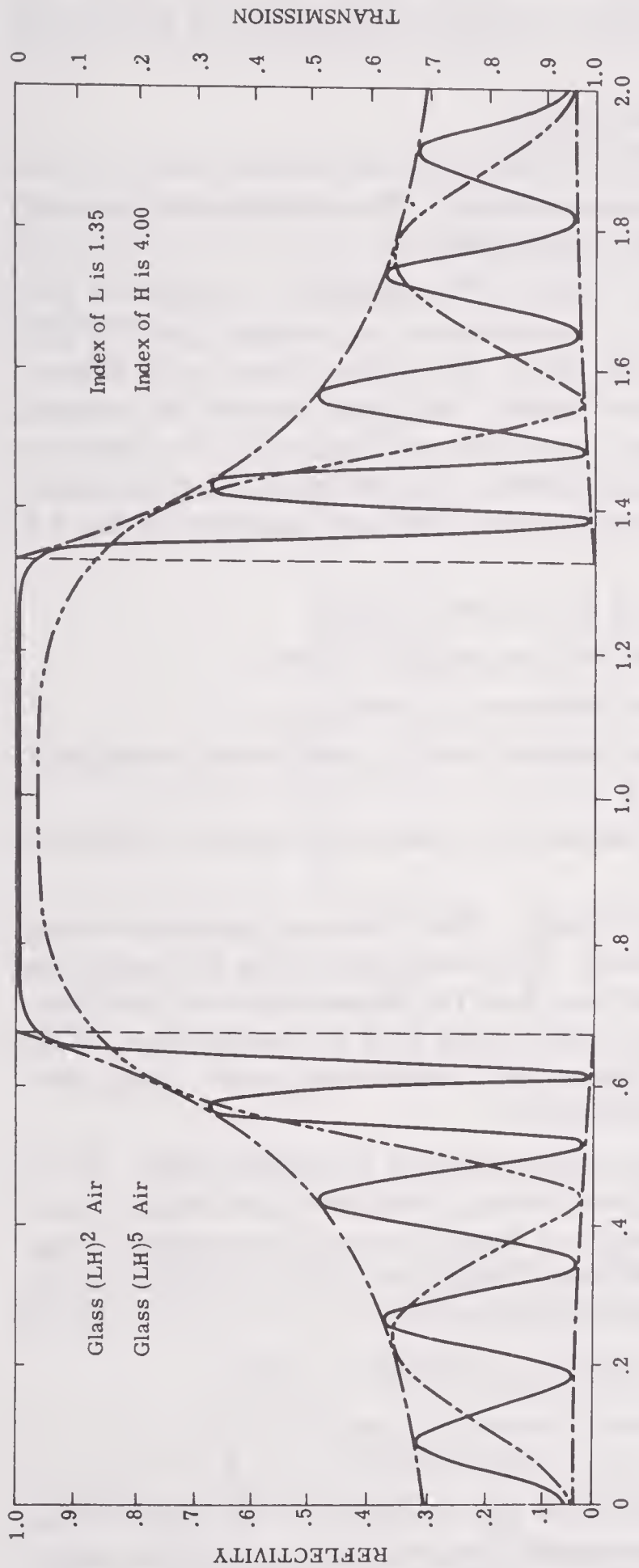


FIG. 7-12. Computed spectral reflectivity of a four-layer (– – –) and a ten-layer (———) quarter-wave stack, and envelope of maximum reflectivity.

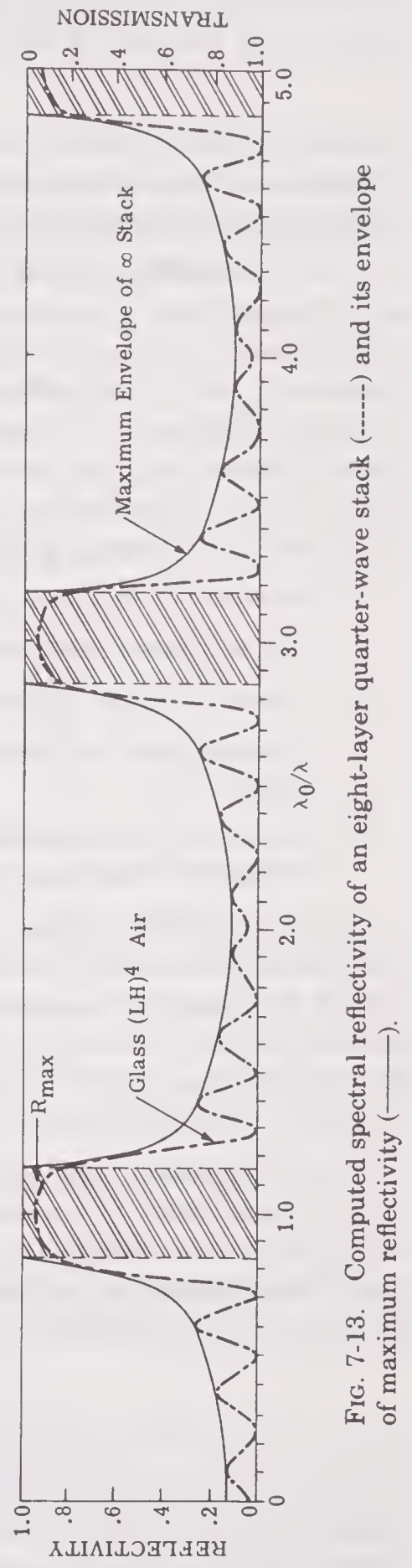


FIG. 7-13. Computed spectral reflectivity of an eight-layer quarter-wave stack (-----) and its envelope of maximum reflectivity (———).

3. The number of oscillations outside the high-reflectance zone increases as the number of layers is increased.
4. Other things being equal, the quarter-wave stack has the highest reflectivity.

Additional layers can be added to: (1) increase the reflection in the stop band, or (2) decrease the reflection in the passband. To increase reflectivity use H layers at both ends of the basic period:

$$HLHLHLH$$

To decrease the reflection, the layer can be replaced by its Herpin equivalent and antireflection coatings designed for the desired wavelength. The stack can also be varied by computer techniques based on variational principles.

7.4.4. Long-Wave Pass Interference Filters. The design of a long-wave pass interference filter is based on Fig. 7-12. Low-reflectance regions are regions of high transmission, and can be designed on that basis. The curves change with different ratios of n_H to n_L and different numbers of layers. The design proceeds by choosing a useful substrate and a design λ_0. Some changes can then be made. The long-wave cutoff is determined by either the substrate absorption or the second-order maximum (Fig. 7-13). Commercially available long-wave pass filters are described in Sec. 7.9. They have the following properties:

1. The slope of the cuton increases with the number of layers.
2. The maximum reflectance increases with the number of layers.
3. The width of the reflectance zone increases as n_H/n_L increases.
4. A higher-order reflectance peak has a sharper cuton but a narrower transmission region ($\Delta\lambda$).
5. Angle shift is minimized by high values for n_H and n_L or by more high-index materials in the basic period.

7.4.5. Short-Wave Pass Interference Filters. The comments applicable to long-wave pass filters (Sec. 7.4.4) also apply here. Short-wave pass filters are usually designed from quarter-wave stacks because these have the longest region of high transmission to the short-wave (high-frequency) side of the high reflectance zone. Then an antireflection coat is applied to the stack in the transmission region. Some commercially available filters are described in Sec. 7.9.

7.4.6. Bandpass Interference Filters. Every filter is a bandpass filter. If the desired pass region is smaller than that obtained by a long-wave pass filter, a short-wave pass can be added; they can be deposited on opposite sides of the substrate. Narrower bandpasses are obtained by interference techniques, similar to that for the Fabry-Perot interferometer. The transmission is given by

$$T = \frac{T_1 T_2}{\left(1 - \sqrt{R_1 R_2}\right)^2}\left[1 + \frac{4\sqrt{R_1 R_2}\,\sin\left(\dfrac{2\pi nt}{\lambda} - \dfrac{\epsilon_1 + \epsilon_2}{2}\right)}{\left(1 - \sqrt{R_1 R_2}\right)^2}\right]$$

where T_1, T_2, R_1, and R_2 are the transmittance and reflectance of the plate coatings (looking from the gap); ϵ_1 and ϵ_2 are the phase shift upon reflection, and nt is the optical path of the gap. The transmission is a sinusoidal function of $1/\lambda$ or σ. The region between adjacent transmission peaks is the free spectral range. In filter language this is the free filter range, or *FFR*. This spacing σ_f is given by

$$\sigma_f = \frac{1}{2t_g}$$

The narrowness of a line is given by the Q:

$$Q = \frac{\lambda_0}{(\Delta\lambda)_{0.5}} = \frac{\sigma_0}{(\Delta\sigma)_{0.5}} = \frac{(R_1R_2)^{1/4}}{1 - \sqrt{R_1R_2}} m\pi$$

If R_1 and R_2 are large and if ϵ_1 and ϵ_2 are constant over $\Delta\lambda$, then the line has a Lorentz shape. Narrowband filters can have the following construction

$$\overbrace{HLH}^{A}\ LL\ \overbrace{HLH}^{B}$$

This can be thought of as two filters separated by a half-wave of low index material. The curve for this filter is shown in Fig. 7-14. These filters are usually combined with blocking filters to isolate the narrow band. Some commercially available filters are described in Sec. 7.9.

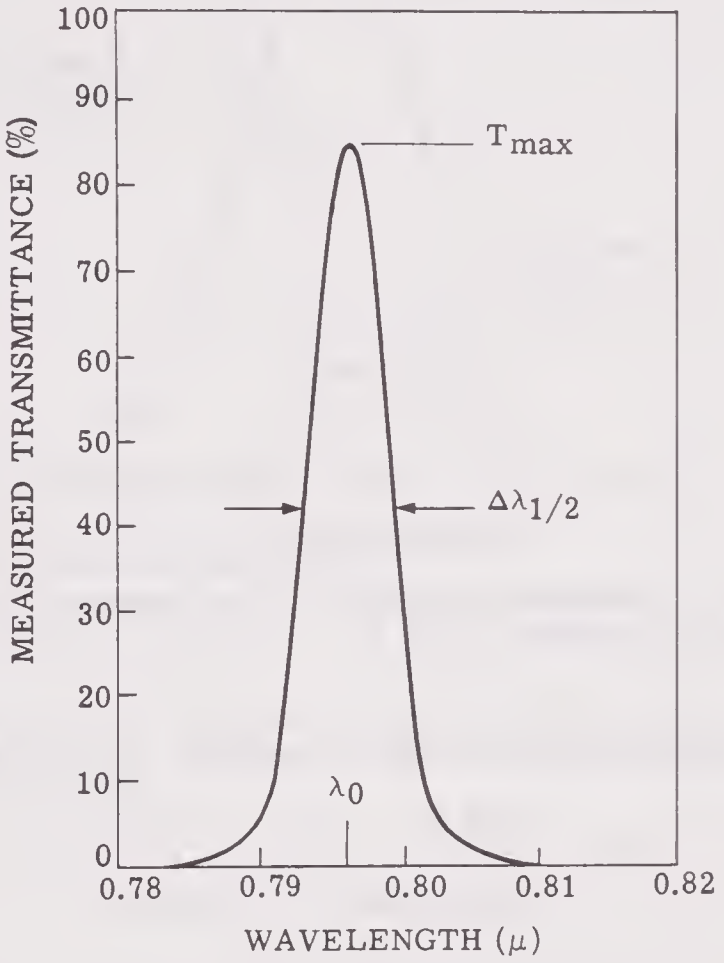


FIG. 7-14. Measured transmission of a narrowband filter.

7.4.7. "Square-Band" Interference Filters. This type of filter—not really a square band—can be designed as a general p:q stack. The design is

$$(HL)^m\ LL\ (HL)^m\ LL \cdots LL\ (HL)^m$$

The filter is generally steeper and has a rippled top. The rejection is also better than the quarter-wave stack. A comparison of the "square" filter with the normal quarter-wave stack is given in Fig. 7-15. These can also be thought of as multiple-wave filters.

7.4.8. The Filter Matrix. The relations between the electric and magnetic fields on two sides of the ith interface are given by

$$\begin{matrix} E \\ H \end{matrix}\Bigg] = M_i \begin{matrix} E' \\ H' \end{matrix}\Bigg]$$

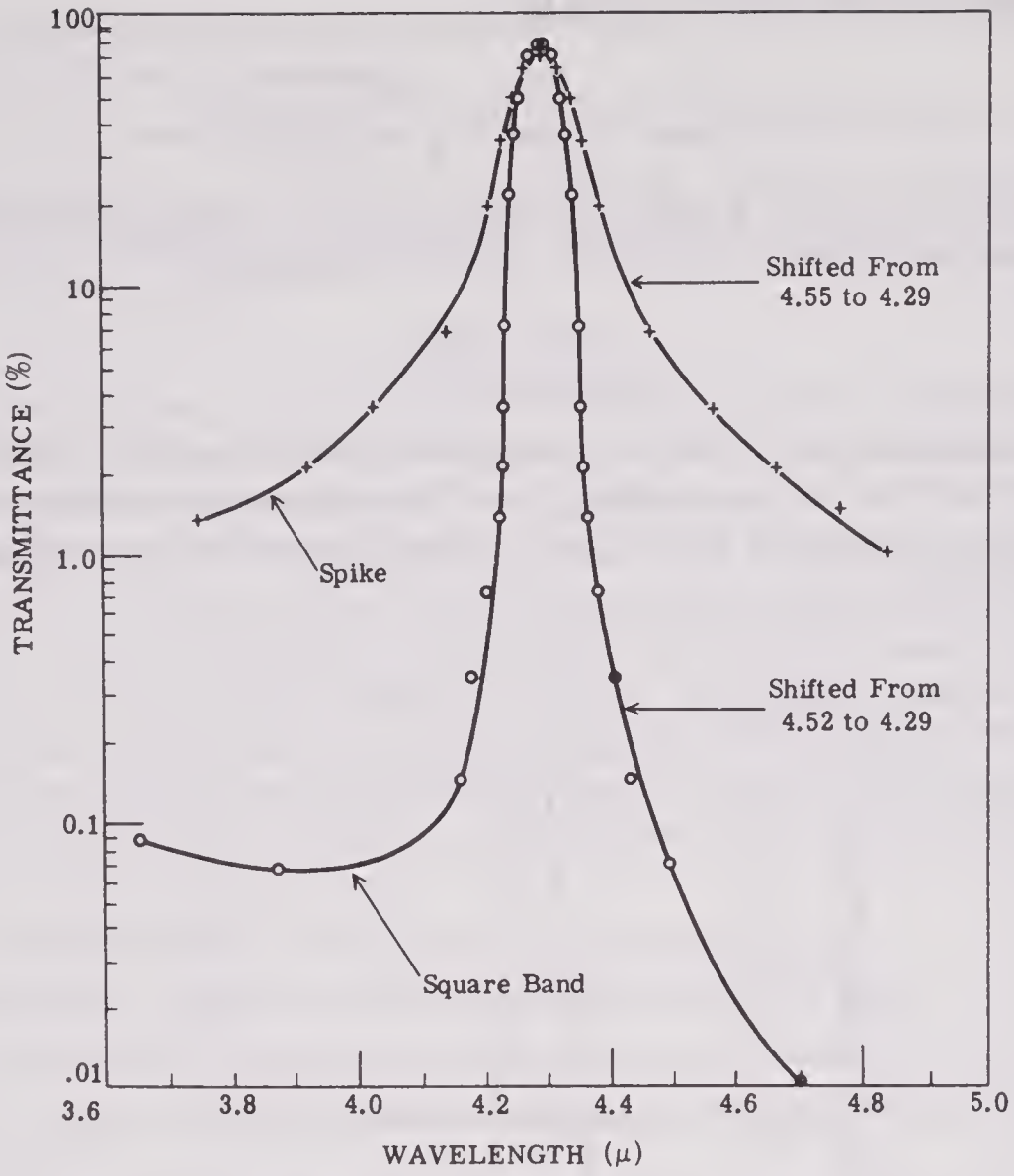


FIG. 7-15. Measured transmittance of two bandpass filters with nominal 2% halfwidth at 4.29 μ.

Here M_i is the characteristic matrix of the ith surface. M_i is given by

$$M_i = \begin{bmatrix} \cos \delta_i & jn_{-1} \sin \delta_i \\ jn_i \sin \delta_i & \cos \delta_i \end{bmatrix}$$

The relation between

$$\begin{matrix} E_0 \\ H_0 \end{matrix} \quad \text{and} \quad \begin{matrix} E_m \\ H_m \end{matrix}$$

is

$$\begin{matrix} E_0 \\ H_0 \end{matrix} = \prod_{i=1}^{m} M_i \begin{matrix} E_m \\ H_m \end{matrix} = M \begin{matrix} E_m \\ H_m \end{matrix} = \begin{bmatrix} a_{11} & ja_{12} \\ ja_{21} & a_{22} \end{bmatrix} \begin{matrix} E_m \\ H_m \end{matrix}$$

The determinant $|M|$ is 1, so that the reflectivity R can be calculated from a knowledge of three elements.

For any periodic layer, the period can be reduced to a fictitious bilayer, and if the period occurs m times, then

$$\begin{matrix} E_0 \\ H_0 \end{matrix} = (M_1 M_2)^m \begin{matrix} 1 \\ n_s \end{matrix}$$

For a symmetrical layer the period can be replaced by a monolayer.

7.4.9. The Herpin Equivalent Layer. The thickness of the layer d_n can be written in terms of the phase δ_n:

$$\begin{aligned}\delta_h &= 360nt/\lambda \quad \text{(degrees)}\\ &= 2\pi nt/\lambda \quad \text{(radians)}\\ &= nt/\lambda \quad \text{(wavelengths)}\\ &= \arccos a_{11}\\ n_h &= a_{21}/(1 - a_{11}^2)^{1/2}\end{aligned}$$

7.4.10. Analogies with Transmission-Line Theory [4]. The matching theorems involving calculation of line admittance, characteristic admittance, reflection coefficient, etc., can be applied to optical multilayers by treating the refractive index as the admittance. Some useful equations are:

For the nth element of an infinite lumped-constant line (s = series; sh = shunt):

$$i_n = Ae^{\gamma n}$$

$$\cosh\gamma = 1 + \frac{Z_s}{2Z_{sh}}$$

$$e^{\gamma} = 1 + \frac{Z_s}{2Z_{sh}} \pm \sqrt{\frac{Z_s}{Z_{sh}} + \left(\frac{Z_s}{2Z_{sh}}\right)^2}$$

$$Z_0 = \pm\sqrt{Z_s Z_{sh} + (Z_s/2)^2}$$

$$\sinh\gamma = Z_0/Z_{sh}$$

For an infinite distributed-parameter line:

$$Z = R + j\omega L \qquad Y = G + j\omega C$$

$$Z_0 = \pm\sqrt{Z/Y} \qquad \gamma = \pm\sqrt{ZY} = Z_0 Y$$

For a line terminated by Z_R:

$$R = \frac{Z_0 - Z_R}{Z_0 + Z_R} = \frac{Y_R - Y_0}{Y_R + Y_0}$$

$$Z(l) = Z_0 \frac{Z_0 \sinh \gamma l + Z_R \cosh \gamma l}{Z_0 \cos \gamma l + Z_R \sinh \gamma l}$$

$$Z_{cl} Z_{op} = Z_0^2$$

where Z_{cl} is the short-circuit impedance and Z_{op} is the open-circuit impedance.

Impedance matching: A section of lossless lumped line can be chosen to join generator and load for maximum power transfer if

$$Z_0^2 = Z_{in} Z_{term}$$

and the length is a quarter wave.

7.4.11. Effects of Angle of Incidence. The transmission band is a function of the angle of incidence. An effective optical thickness can be used.

$$(nt)_{eff} = nt \cos\theta$$

where θ = angle of incidence. This technique can be used for a few layers, but since the angle of incidence for each layer is a function of the original angle and all the

preceding layers, the technique is cumbersome. Substituted into the matrix formulation, however, the effective optical thickness is again useful. For nonnormal incidence the filter also becomes polarizing; for simple layers the standard equations for reflectance are useful, or they can be put in matrix form for iteration. Angle effects can be minimized by the use of higher index or using more material of a higher index in a layer. The effective thickness is shown in Fig. 7-16.

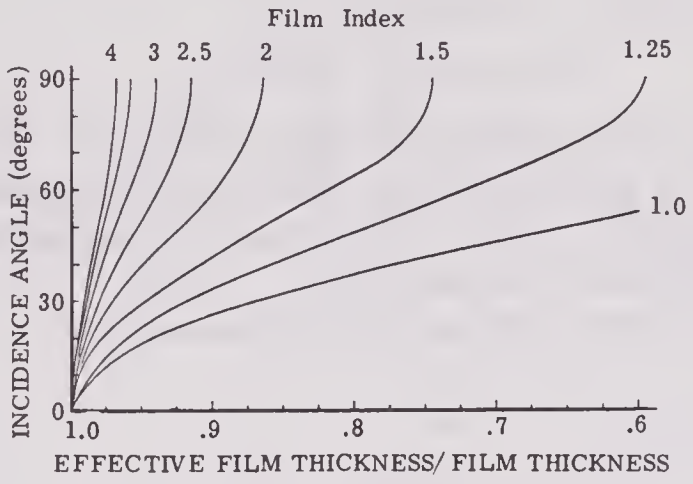


FIG. 7-16. Effective film thickness [5].

7.4.12. Effects of Temperature. Cooling a filter changes both the actual thickness of layers and the refractive index. Thus the filter will change its center wavelength as the optical path changes:

$$\frac{\Delta nt}{\Delta T} = n\frac{\Delta t}{\Delta T} + t\frac{\Delta n}{\Delta T}$$

These terms are calculable from data of thermal expansion and refractive index change with temperature given in Chapter 8.

7.4.13. Substrates and Films. Table 7-5 is a list of the commonly used substrate materials. Chapter 8 gives the physical data for these materials. Sometimes it will be necessary to extrapolate the data to thinner samples.

TABLE 7-5. COMMONLY USED SUBSTRATE MATERIALS

Material	*Refractive Index*	*Transmission Range* (μ)
Irtran 1	1.38-1.31	1-8.5
Lithium fluoride	1.39-1.30	0.2-9
Calcium fluoride (also as Irtran 3)	1.44-1.32	0.2-12
Vycor	1.46	0.25-3.5
Fused quartz	1.48-1.41	0.2-4.5
Barium fluoride	1.51-1.41	0.2-15
Glass	1.70-1.51	0.32-2.5
Magnesium oxide	1.77-1.62	0.3-8.5
Sapphire	1.83-1.59	0.2-6.5
Irtran 2	2.29-2.15	1-14.5
Irtran 4	~2.4	2.0-24
Arsenic trisulfide glass	2.66-2.37	0.6-11
Silicon	3.50-3.42	1.2-15
Germanium	4.10-4.00	1.8-23

TABLE 7-6. COMMONLY USED FILM MATERIALS. [6]

Material	*Refractive Index*	*Range of Transparency*[9]		*Comments*
		From	*To*	
Cryolite	1.35	<200 mμ	10 μ	1
Chiolite	1.35	<200 mμ	10 μ	1
Magnesium fluoride	1.38	230 mμ	5 μ	2,3
Thorium fluoride	1.45	<200 mμ	10 μ	–
Cerium fluoride	1.62	300 mμ	> 5 μ	4
Silicon monoxide	1.45 to 1.90	350 mμ	8 μ	5
Sodium chloride	1.54	180 mμ	>15 μ	6
Zirconium dioxide	2.10	300 mμ	> 7 μ	2
Zinc sulfide	2.30	400 mμ	14 μ	7
Titanium dioxide	2.40 to 2.90	400 mμ	> 7 μ	8
Cerium dioxide	2.30	400 mμ	5 μ	2,3
Silicon	3.50	900 mμ	8 μ	–
Germanium	3.80 to 4.20	1400 mμ	>20 μ	–
Lead telluride	5.10	3900 mμ	>20 μ	–

Notes:

1. Both materials are sodium-aluminum fluroide compounds, but differ in the ratio of Na to Al and have different crystal structure. Chiolite is preferable in the infrared, because it has less stress than cryolite.
2. These materials are hard and durable, especially when evaporated onto a hot substrate.
3. The long wavelength is limited by the fact that, when the optical thickness of the film is a quarter-wave at 5 μ, the film cracks because of the mechanical stress.
4. Other fluorides and oxides of rare earths have refractive indices in this range from 1.60 to 2.0.
5. The refractive index of SiOx (called silicon monoxide) can vary from 1.45 to 1.90 depending upon the partial pressure of oxygen during the evaporation. Films with a refractive index of 1.75 and higher absorb at wavelengths below 500 mμ.
6. Sodium chloride is used in interference filters out to a wavelength of 20 μ. It has very little stress.
7. The refractive index of zinc sulfide is dispersive.
8. The refractive index of TiO_2 rises sharply in the blue spectral region.
9. The range of transparency is for a film of quarter-wave optical thickness at this wavelength. These values are approximate and also depend quite markedly upon the conditions in the vacuum during the evaporation of the film.

Table 7-6 is a list of commonly used film materials [6]. The data for thin evaporated films are often different from those of Chapter 8, which are for solid samples. Since the values vary with deposition conditions, only representative numbers are given in Table 7-6.

7.5. Christiansen Filters [7,8,9]

These filters are made of small, closely packed particles of an infrared-transparent substance suspended in a liquid or a gas. The optical properties of the materials are so chosen that the indices of refraction of the particles and the suspending medium are the same at the wavelength that is to be transmitted. The $dn/d\lambda$ values of the liquid and the solid particles are chosen to be as widely different as possible. Thus, as the wavelength is progressively increased or decreased from the wavelength at which equality of the indices occurs, the difference in index between the particles and the suspending medium increases rapidly.

One form of Christiansen filter for the infrared is obtained by using quartz particles in air. Figure 7-17 shows the dispersion curve of ordinary quartz. The refractive index is unity at 7.4 μ. At this wavelength, therefore, the quartz particles have the same refractive index as air, and high transmission occurs.

Table 7-7 lists other materials which when suspended in air, can be used as Christiansen filters. Also shown are the wavelengths at which maximum transmission occurs. These are designated as Christiansen wavelengths.

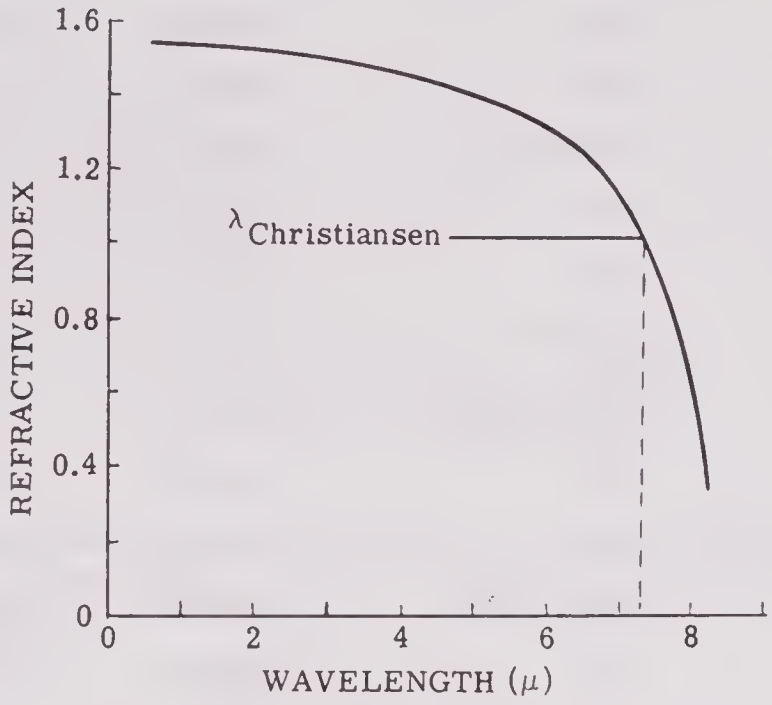


FIG. 7-17. Dispersion curve of quartz, showing Christiansen wavelength [7].

TABLE 7-7. CHRISTIANSEN WAVELENGTHS OF SELECTED MATERIALS [7]

Crystal in Air	*Christiansen Wavelength* (μ)
Quartz	7.3
LiF	11.2
MgO	12.2
NaCl	32
NaBr	37
NaI	49
KBr	52
KI	64
RbI	73
TlI	90

Figure 7-18 shows the effect of quartz powder in a medium of pure CCl_4 and CS_2 as compared with quartz in air. Because values of dn/dT are relatively high, these filters are sensitive to temperature fluctuations. By the same token, in a controlled environment the center of the passband can be varied by changing the temperature.

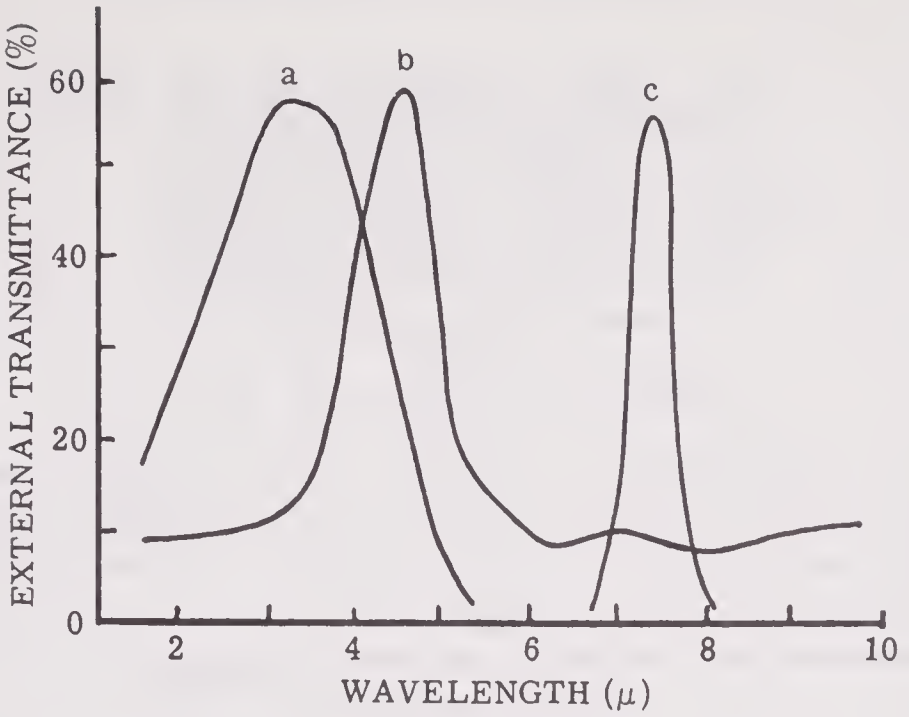


FIG. 7-18. Position of the Christiansen peak for quartz power in liquids. (*a*) Quartz in a 50% by volume mixture of CS_2 and CCl_4. (*b*) Quartz in pure CCl_4. (*c*) Quartz in air [10].

7.6. Selective Reflection Filters [11]

Selective reflection filters are made of crystalline materials that show selective reflection at certain wavelengths. These filters are useful to about 200 μ.

In practice, radiation from a suitable source is collimated and directed at the surface of a polished crystal whose residual ray occurs at the wavelength to be selectively reflected. After three or four successive reflections from similar crystal plates, only the residual ray is present with any appreciable intensity, the other wavelengths having been attenuated by a factor of several thousand. Figures 7-19 and 7-20 indicate the wavelengths of the residual rays of a number of materials.

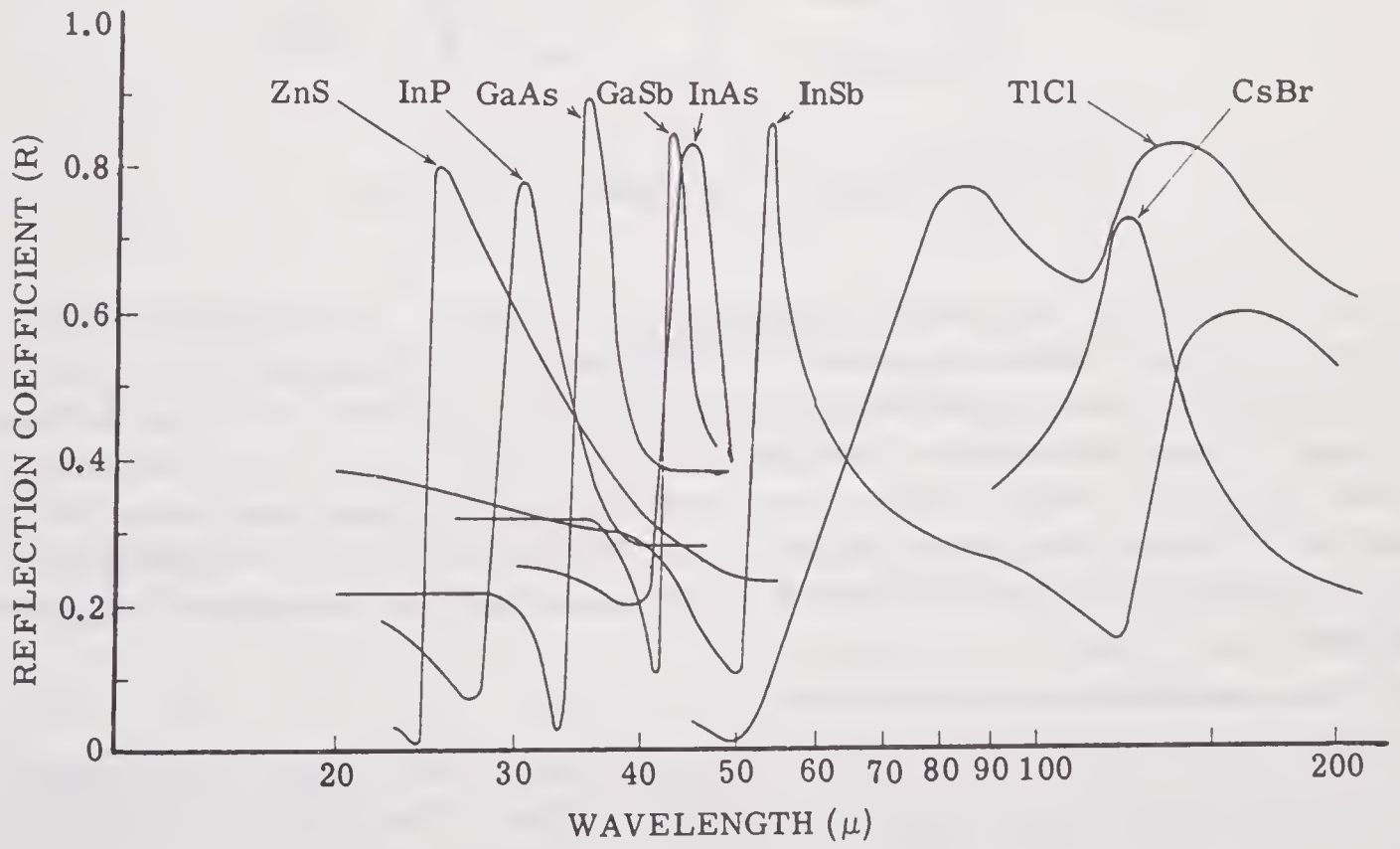


FIG. 7-19. Reststrahlen (residual ray) frequencies of alkali halide crystals.

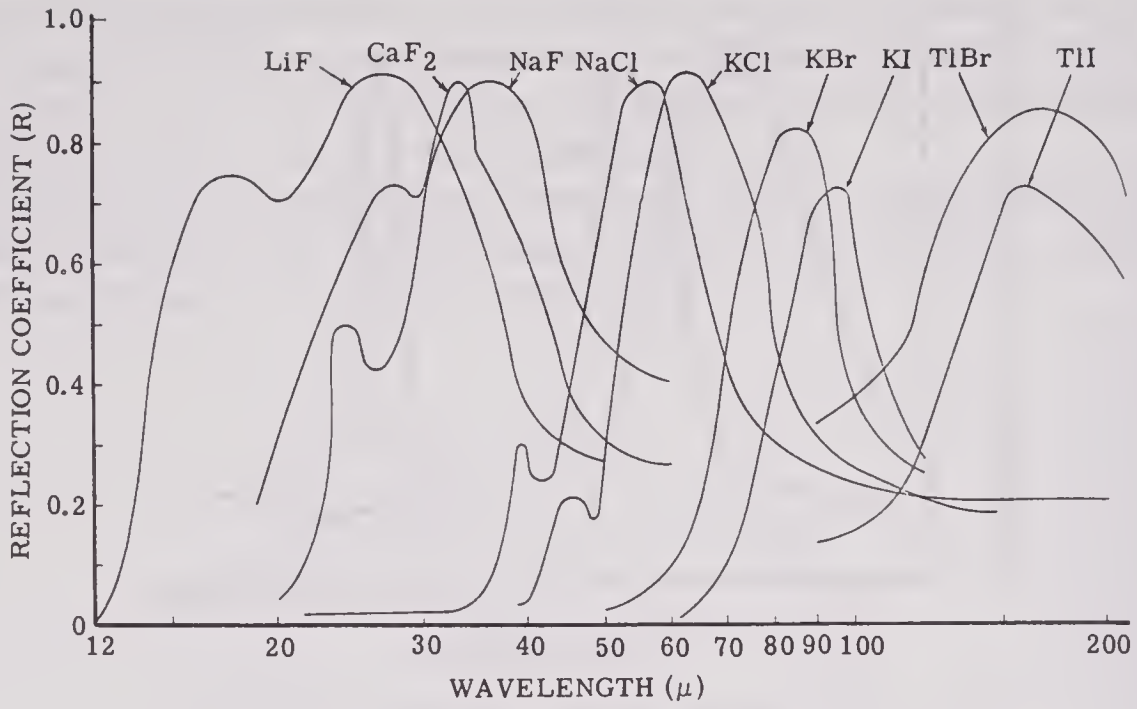


FIG. 7-20. Reststrahlen (residual ray) frequencies of polar crystals.

7.7. Selective Refraction Filters

Filtering by selective refraction depends upon the $dn/d\lambda$ of a lens material. Radiation of different wavelengths will be focused at different points along the optical axis if the lens is used in a wavelength region where it possesses high dispersion. This technique is particularly useful when the lens is used near an absorption band, because the refractive index will be considerably different on opposite sides of the band. This method of focal isolation or selective refraction is illustrated in Fig. 7-21.

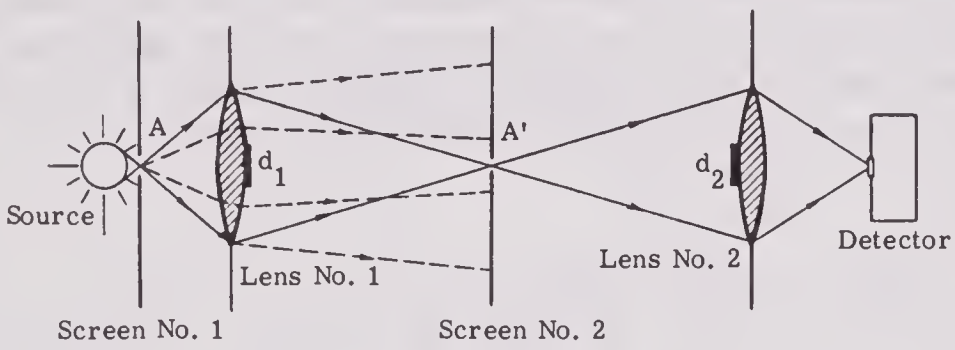


FIG. 7-21. Filtering by the method of focal isolation.

Two quartz lenses are usually employed in this method. The refractive index of quartz in the near infrared is about 1.5, and about 2.15 in the range 60 μ to 100 μ. (Quartz absorbs in the region around 9μ.) In Fig. 7-21, radiation in the far infrared is passed through the aperture, brought to a focus at aperture A', and transmitted. Visible and near-infrared radiation, being deviated less, impinges upon screen No. 2. Lens No. 2 focuses the desired radiation upon the detector. The two opaque discs, d_1 and d_2, obscure the paraxial zone of the lenses and prevent transmission of direct radiation.

7.8. Polarization Interference Filters [11]

The polarization interference filter, sometimes called a Lyot-Ohman (or birefringent) filter, isolates a spectral band only a few angstroms wide. These filters are constructed of alternate plates of polarizers and birefringent crystals (*e.g.*, quartz) as shown in Fig. 7-22.

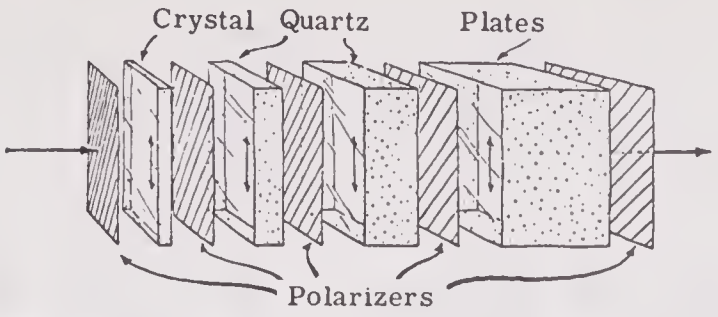


FIG. 7-22. Birefringent filter [11].

The birefringent crystal and quartz plates are cut with their optical axes parallel to the large faces. The axes of the polarizers are oriented at 45° to the quartz optical axes. Linearly polarized light, incident on the first quartz plate at 45°, would have its plane of polarization rotated by 90° if the plate were a half-wave plate (or if the optical path difference between the ordinary and extraordinary rays were any odd multiple of half-waves). If the plane of polarization is rotated at 90°, the radiation will not be transmitted by the second polarizer. Since the phase difference introduced between the ordinary and extraordinary rays depends on wavelength as well as on the thickness of the quartz plate, the same plate may be a 5/2-wave plate for one wavelength, a 7/2-wave plate for another wavelength, and 9/2-wave plate for still another wavelength. Each of these wavelengths will be blocked by the polarizer following the plates, whereas those wavelengths for which the difference in optical path between the two polarizations is an even number of half waves will be completely transmitted.

Figure 7-23(*a*) illustrates the transmission of the first quartz plate and its polarizers. If each quartz plate is made twice as thick as the preceding one, it will have twice as many transmission maxima and minima in a given wavelength interval. The transmission curves for the second, third, and fourth plates are illustrated in Fig. 7-23 (*b*), (*c*), and (*d*). The transmission of the entire filter is the result of all these transmission curves and is shown in Fig. 7-23(*e*).

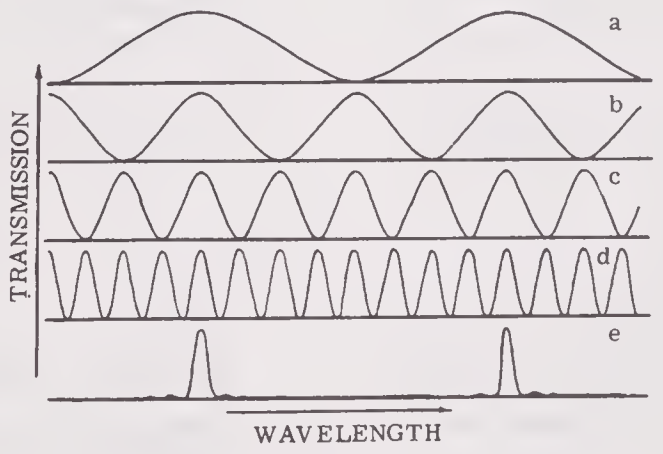
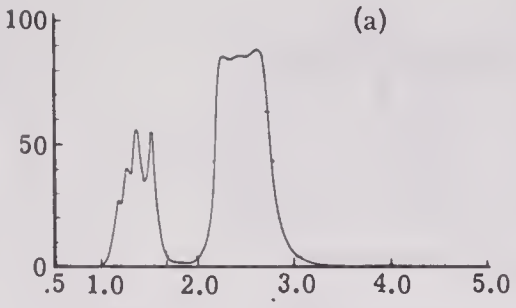


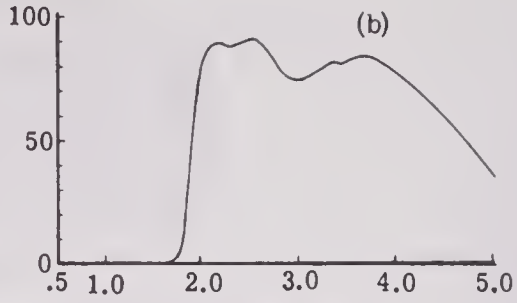
FIG. 7-23. Transmission of Lyot filter and its components [11].

7.9. Commercially Available Filters

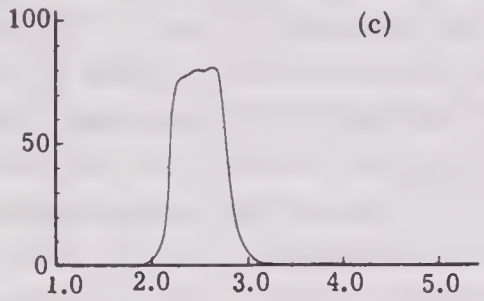
Figure 7-24 presents transmittance curves of typical Bausch & Lomb filters. Figure 7-25 shows transmittance curves of interference filters available from Eastman Kodak. Figure 7-26 shows transmittance curves of interference filters available from Infrared Industries, Inc. Figure 7-27 shows transmittance curves of filters available from Optical Coating Laboratory, Inc. Other infrared filter manufacturers include Optics Technology, Inc., Fish Schurman for Schott Glasswerke, Baird-Atomic (mainly for visible and ultraviolet), Farrand Optical Company, Inc.; a number of other companies are listed in the Optical Industry catalog.



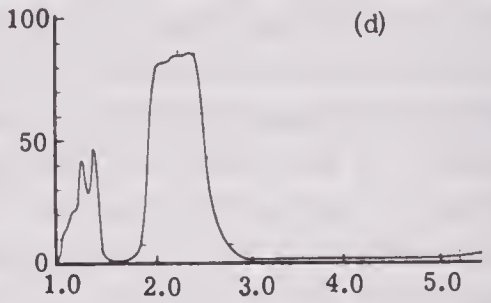
Bandpass filter (cover glass substrate) $\lambda_0 = 2.45\ \mu$



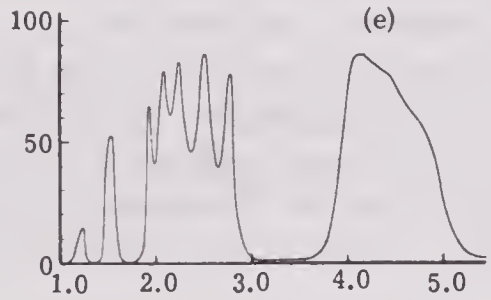
Long-wave pass filter (cover glass substrate) $\lambda_0 = 1.8\ \mu$



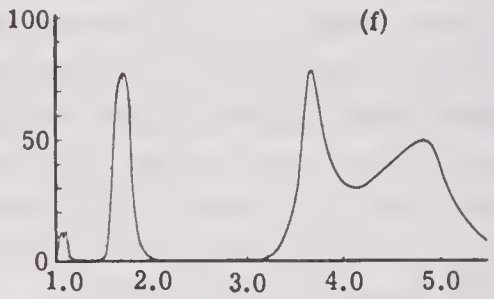
Bandpass filter in series with auxiliary long-wave pass filter (cover glass substrate) $\lambda_0 = 2.48\ \mu$



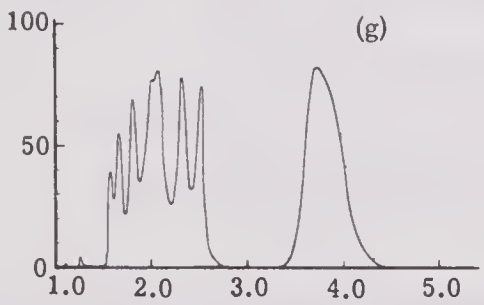
Bandpass filter (cover glass substrate) $\lambda_0 = 2.26\ \mu$



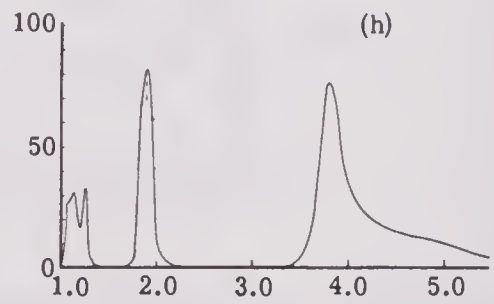
Bandpass filter (cover glass substrate) $\lambda_0 = 4.42\ \mu$



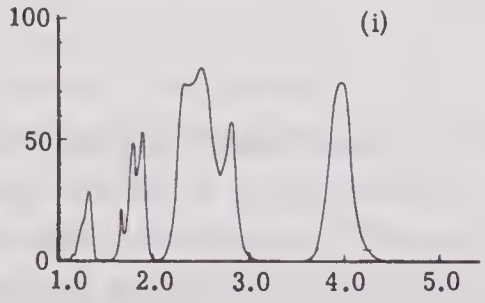
Bandpass filter (cover glass substrate) $\lambda_0 = 1.68\ \mu$



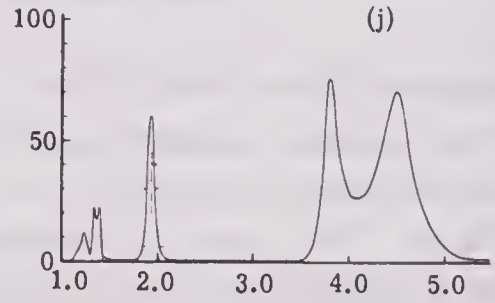
Bandpass filter (cover glass substrate) $\lambda_0 = 3.82\ \mu$



Bandpass filter (cover glass substrate) $\lambda_0 = 1.89\ \mu$

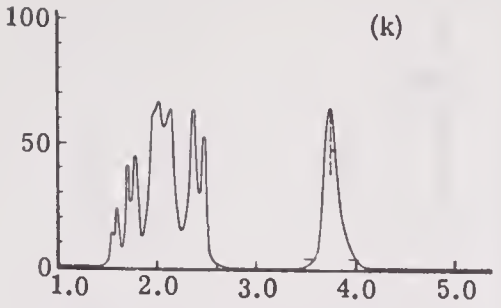


Bandpass filter (cover glass substrate) $\lambda_0 = 3.95\ \mu$

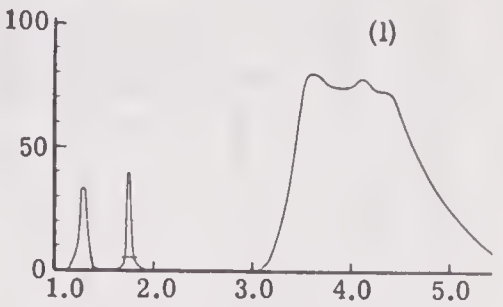


Bandpass filter (cover glass substrate) $\lambda_0 = 1.94\ \mu$

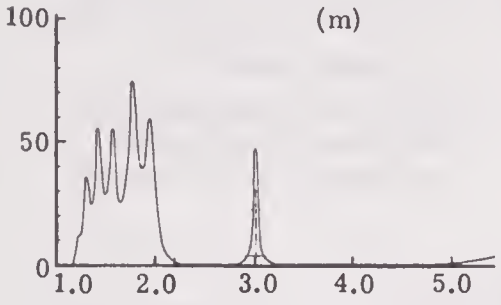
FIG. 7-24. Transmittance of Bausch & Lomb infrared interference filters [4].



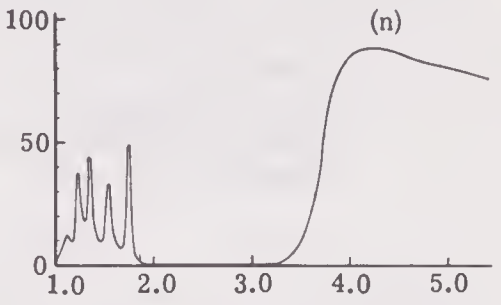
Bandpass filter (cover glass substrate) $\lambda_0 = 3.74\ \mu$



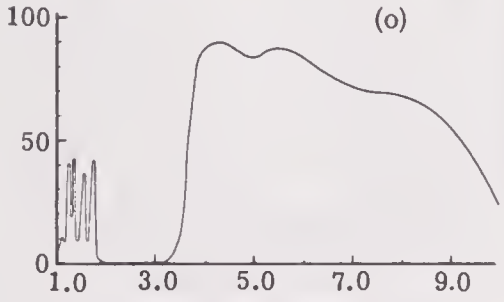
Bandpass filter (cover glass substrate) $\lambda_0 = 0.175\ \mu$



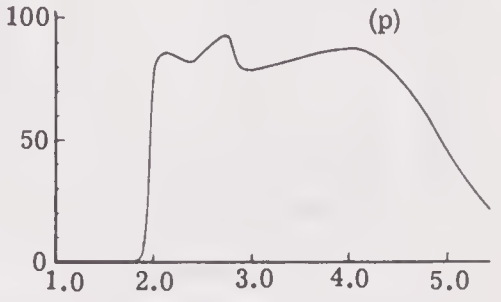
Bandpass filter (cover glass substrate) $\lambda_0 = 3.01\ \mu$



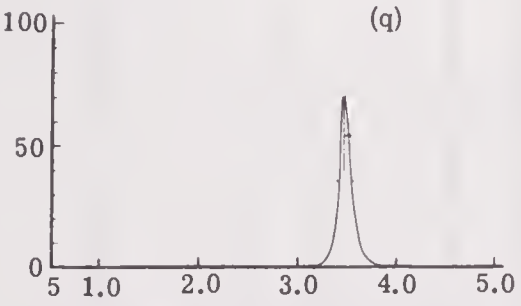
Long-wave pass filter (magnesium oxide substrate) $\lambda_0 = 3.44\ \mu$



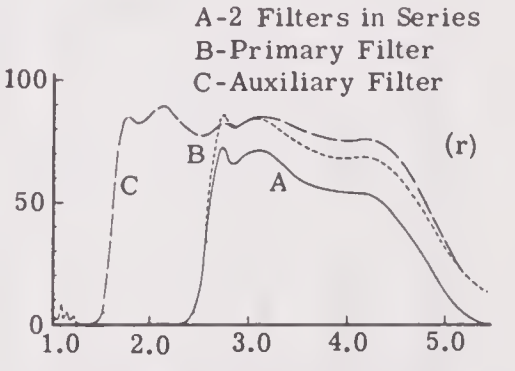
Long-wave pass filter (calcium fluoride substrate) $\lambda_0 = 3.4\ \mu$



Long-wave pass filter (cover glass substrate) $\lambda_0 = 1.9\ \mu$

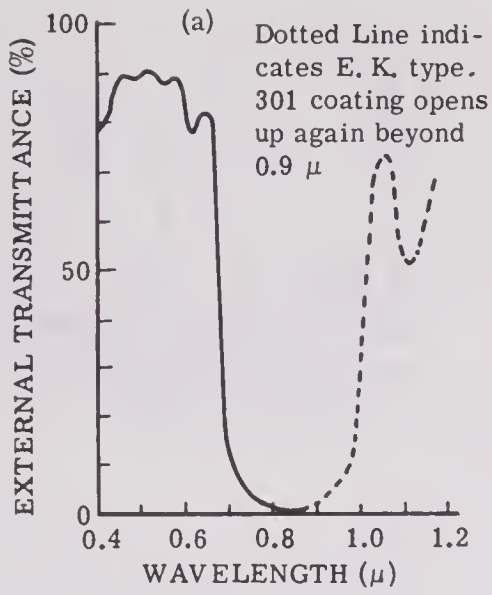


Bandpass filter ($\lambda_0 = 3.45\ \mu$) in series with long-wave pass filter ($\lambda_0 = 2.6\ \mu$)

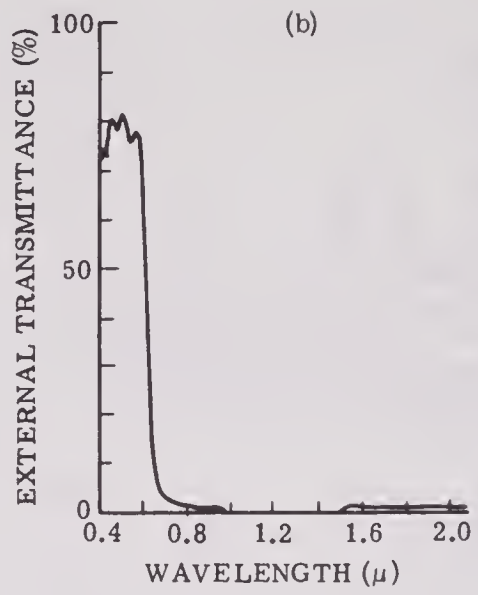


Long-wave pass filter ($\lambda_0 = 2.44\ \mu$) in series with long-wave pass filter ($\lambda_0 = 1.5\ \mu$); cover glass substrate

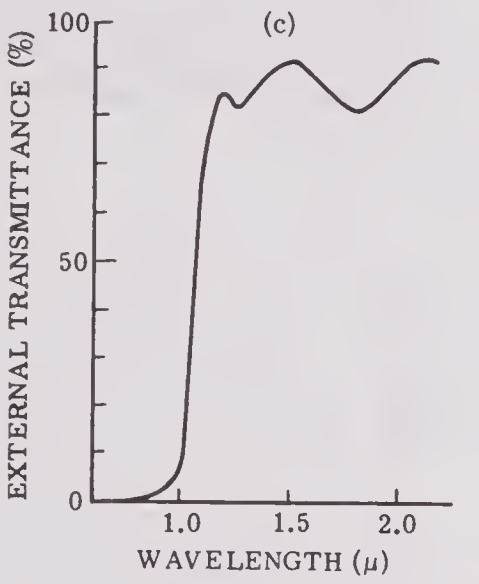
FIG. 7-24 (*Continued*). Transmittance of Bausch & Lomb infrared interference filters [4].



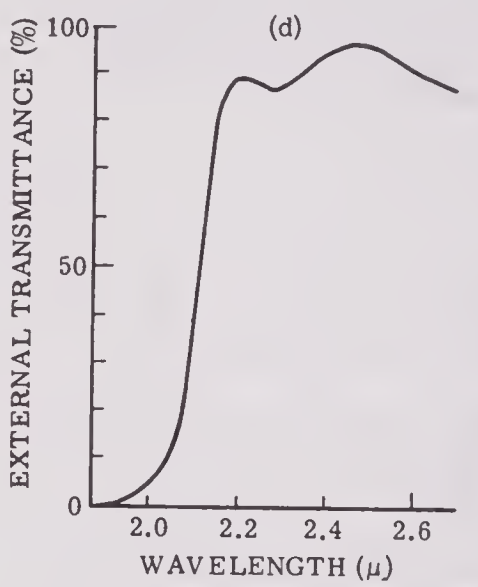
Short-wave pass filter (plate glass substrate with Kodak No. 301 coating)



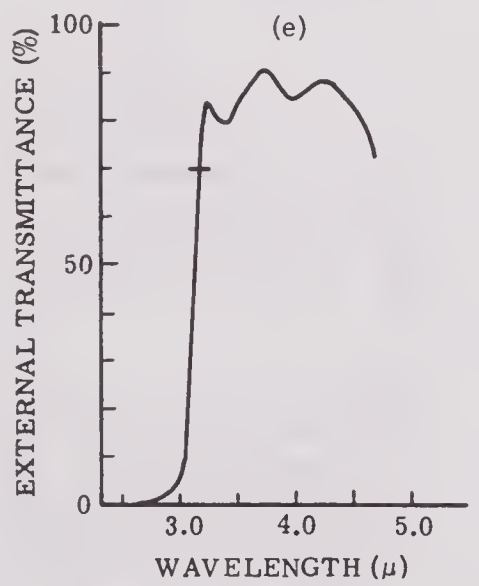
Short-wave pass filter (heat-absorbing glass substrate with Kodak No. 301 coating)



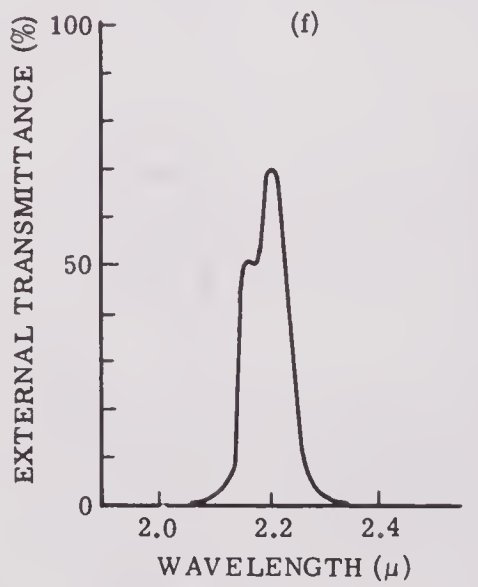
Long-wave pass filter; $\lambda_0 = 1.0\,\mu$



Long-wave pass filter; $\lambda_0 = 2.0\,\mu$

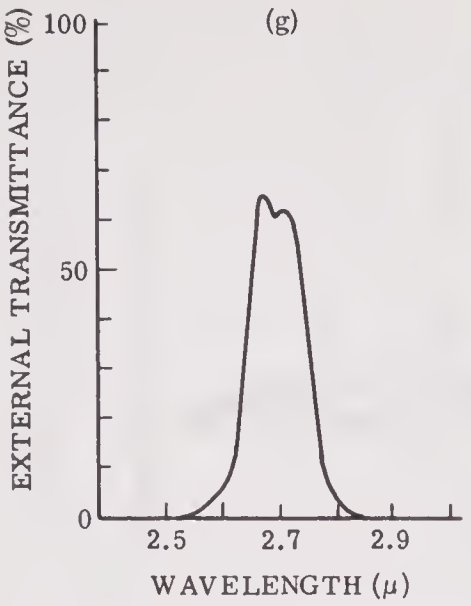


Long-wave pass filter; $\lambda_0 = 3.0\,\mu$

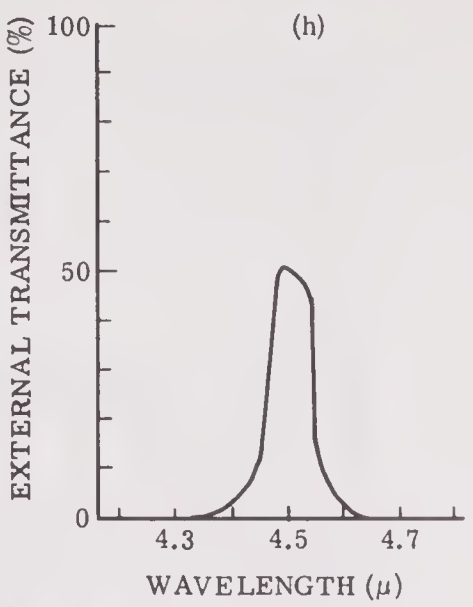


Bandpass filter; $\lambda_0 = 2.2\,\mu$

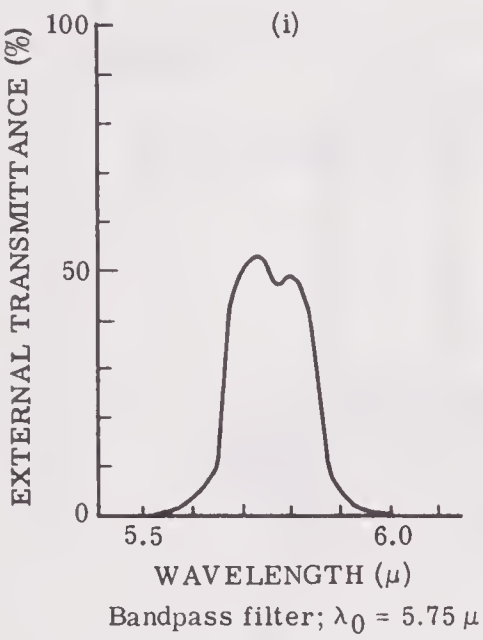
FIG. 7-25. Transmittance of Eastman Kodak infrared interference filters [1].



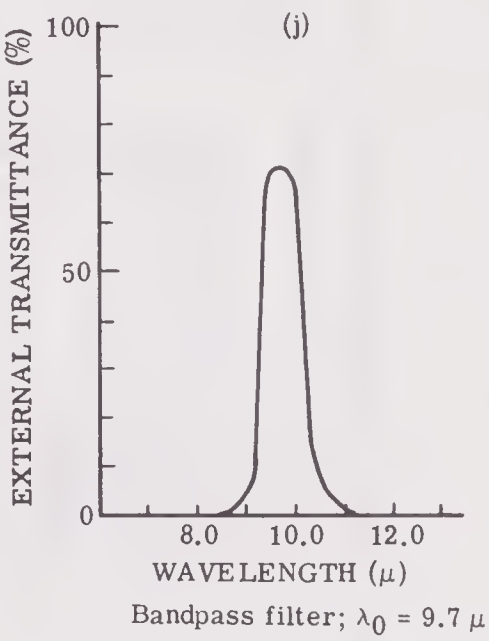
Bandpass filter; $\lambda_0 = 2.7\ \mu$



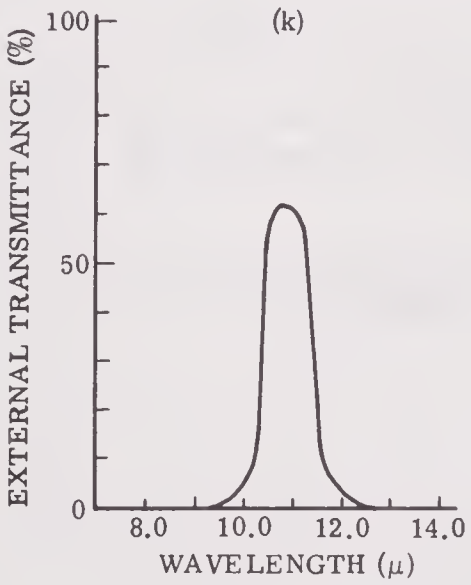
Bandpass filter; $\lambda_0 = 4.5\ \mu$



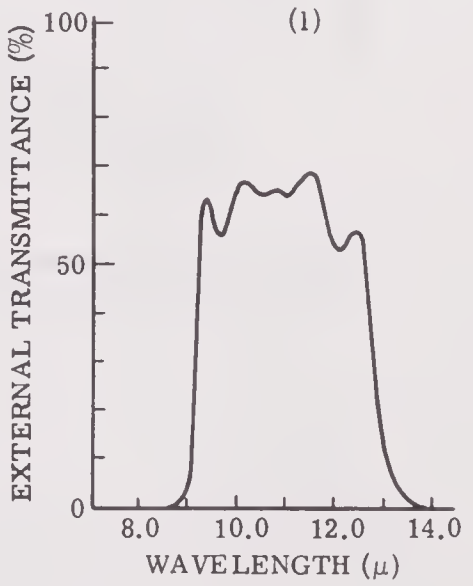
Bandpass filter; $\lambda_0 = 5.75\ \mu$



Bandpass filter; $\lambda_0 = 9.7\ \mu$

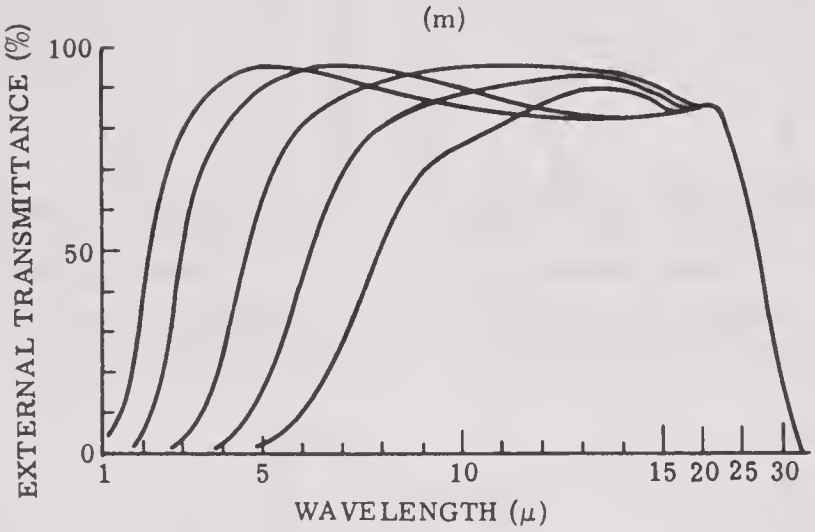


Bandpass filter; $\lambda_0 = 10.9\ \mu$

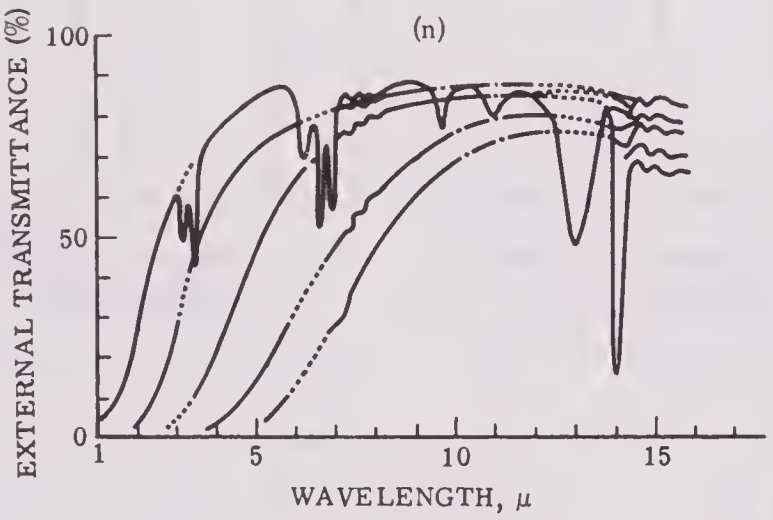


Bandpass filter; $\lambda_0 = 11.0\ \mu$

FIG. 7-25 (*Continued*). Transmittance of Eastman Kodak infrared interference filters [1].

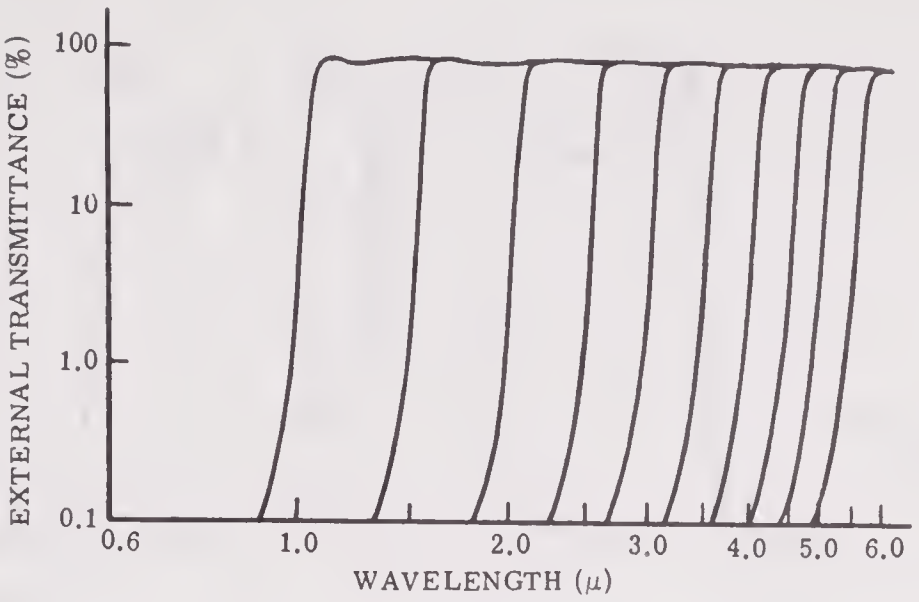


Far-infrared, long-wave pass filters (silver chloride substrate)

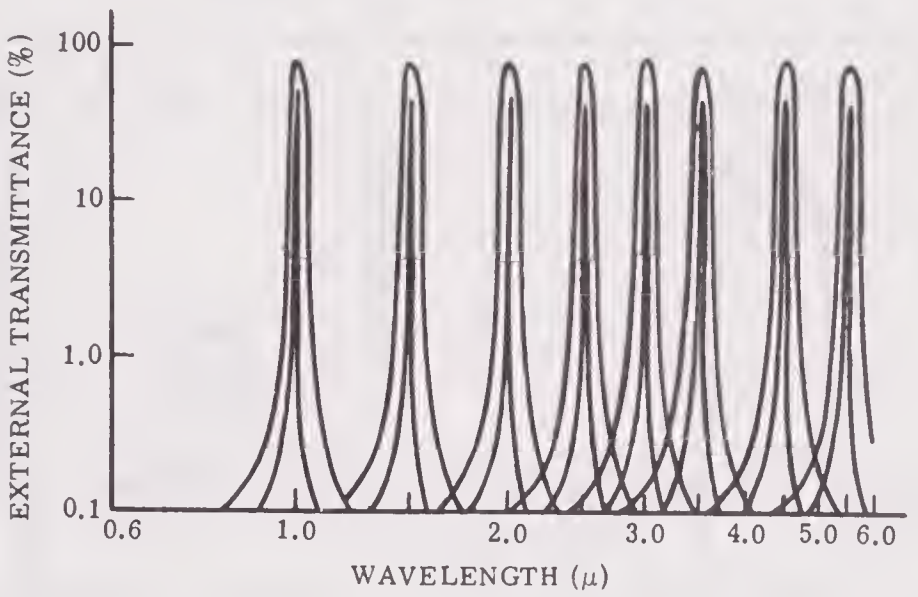


Far-infrared, long-wave pass filters (silver chloride substrate plus polystyrene protective coating)

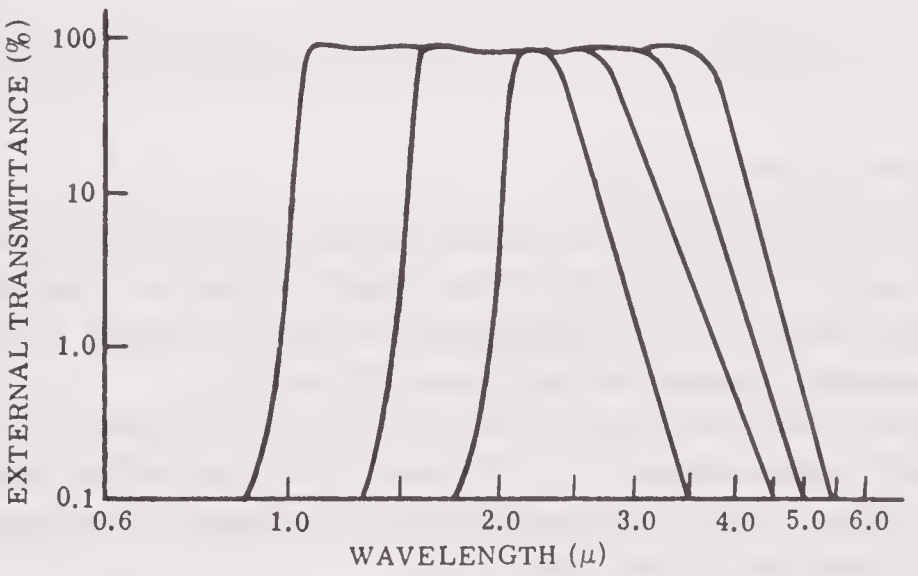
FIG. 7-25 (*Continued*). Transmittance of Eastman Kodak infrared interference filters [1].



(a)



(b)



(c)

FIG. 7-26. Transmittance of Infrared Industries interference (Infratron) filters. (*a*) Long-wave pass; (*b*) spike; (*c*) long-wave pass combined with detector [12].

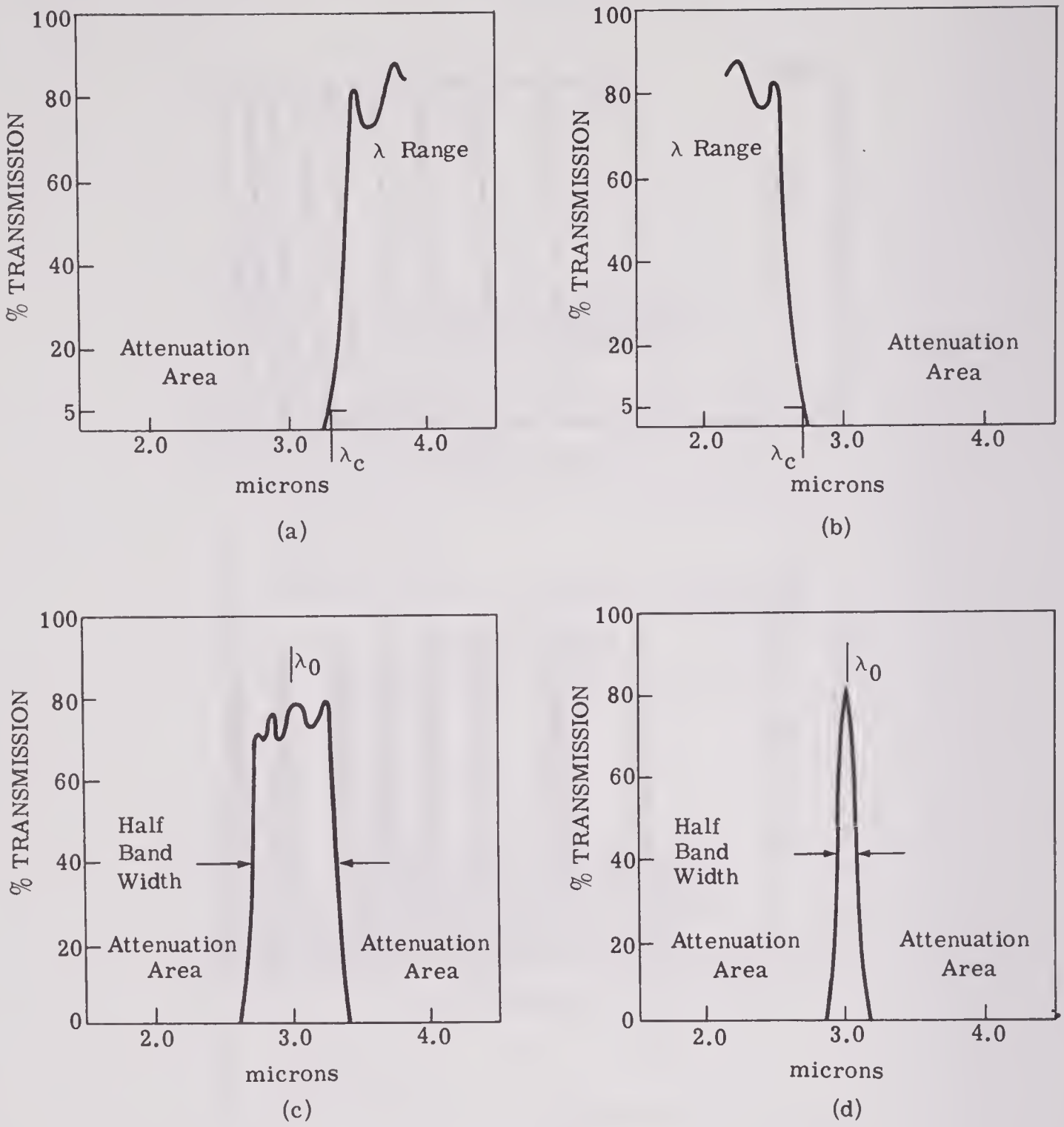


FIG. 7-27. Transmittance of Optical Coating Labs infrared interference filters.

7.10. Absorption Filters [7,12,13]

Many materials in solid, liquid, or gaseous state, including those discussed in Chapter 8, can be used as selective absorption filters in various regions of the infrared spectrum. These filters have high transmission above or below a certain wavelength where high absorption produces a sharp *cutoff* or *cuton*.

Long-wave pass filters in the near infrared are normally made of plastic materials containing dyes, colored glass, or sublimated phthalocyanines upon glass. Other long-wave pass filters consist of glass coated with plastic dye solutions. Figure 7-28 shows the characteristics of these filters [14].

Figure 7-29 shows the variety of cuton wavelengths that are obtained with various semiconductors. By proper doping the location of the absorption limit can be moved, although this reduces the gradient. A similar effect can also be obtained by the use of mixed crystals, Fig. 7-33, although no such filters are available commercially at present. An excellent reference on mixed-crystal semiconductors is R. H. Bube [16].

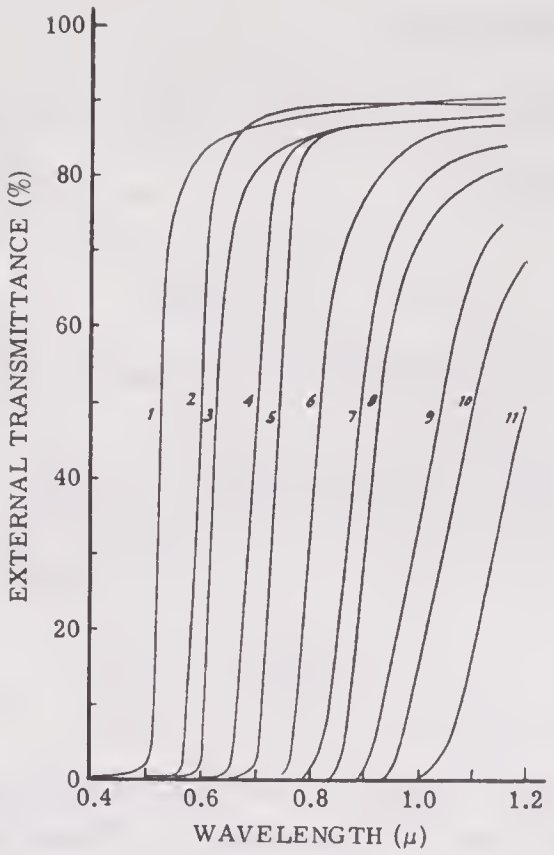


FIG. 7-28. Near infrared dyed-plastic filter characteristics [14].

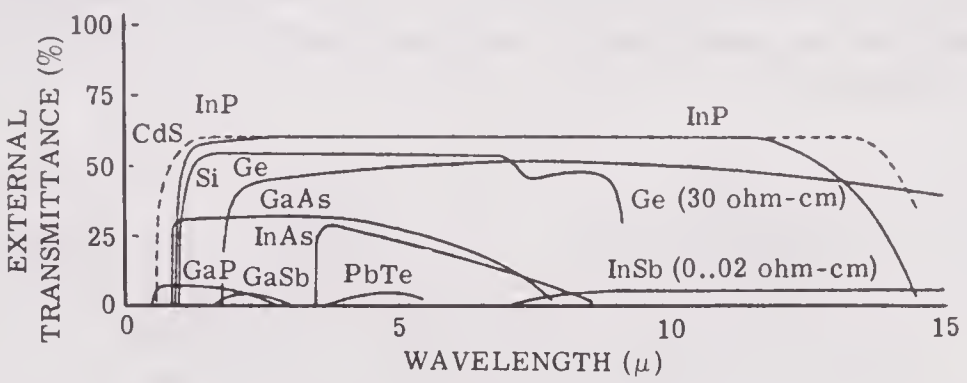


FIG. 7-29. Transmission of selected semiconductors [15].

7.11. Prisms

Prisms are used principally for deviating light or dispersing it. Any prism does both, but it can be arranged to maximize one but minimize the other. Figure 7-30 illustrates the geometry and defines the symbols.

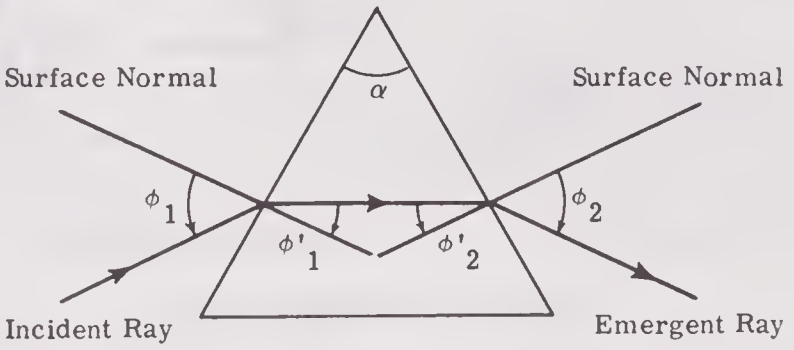


FIG. 7-30. Prism geometry [17].

The angular magnification M_θ is

$$M_\theta = \frac{-\cos\theta_1 \cos\theta_2'}{\cos\theta_2 \cos\theta_1'}$$

The total deviation is

$$\delta = \theta_1 + \theta_2 - \alpha$$

The deviation can also be expressed by

$$\sin \theta_1 = \sin \alpha \sqrt{(n_2^2/n_1^2) - \sin^2\theta_2} - \cos \alpha \sin \theta_2$$

Minimum deviation is given by

$$n_2 = n_1 \frac{\sin \frac{1}{2}(\delta_{min} + \alpha)}{\sin \alpha/2}$$

The change in deviation with wavelength, or the dispersion is

$$\frac{d\delta}{d\lambda} = \frac{dn_2}{d\lambda} \frac{2 \sin \alpha/2}{\sqrt{1 - (n_2/n_1)^2 \sin^2\alpha/2}}$$

The resolving power is

$$\frac{\lambda}{d\lambda} = a \frac{d\theta}{dn} \frac{dn}{d\lambda} = t \frac{dn}{d\lambda}$$

7.11.1. Dispersing Prisms. Figure 7-31 illustrates examples of various types of dispersing prisms, commonly used in spectroscopy. These are the constant deviation type. The Wadsworth prism, operating at minimum deviation, is used extensively in infrared monochromators, where for a constant angle between the collimator and the telescope, astigmatism-free images can be obtained at various wavelengths by rotating the prism. The Littrow prism reflects the light directly back along the direction from which it came [18,19].

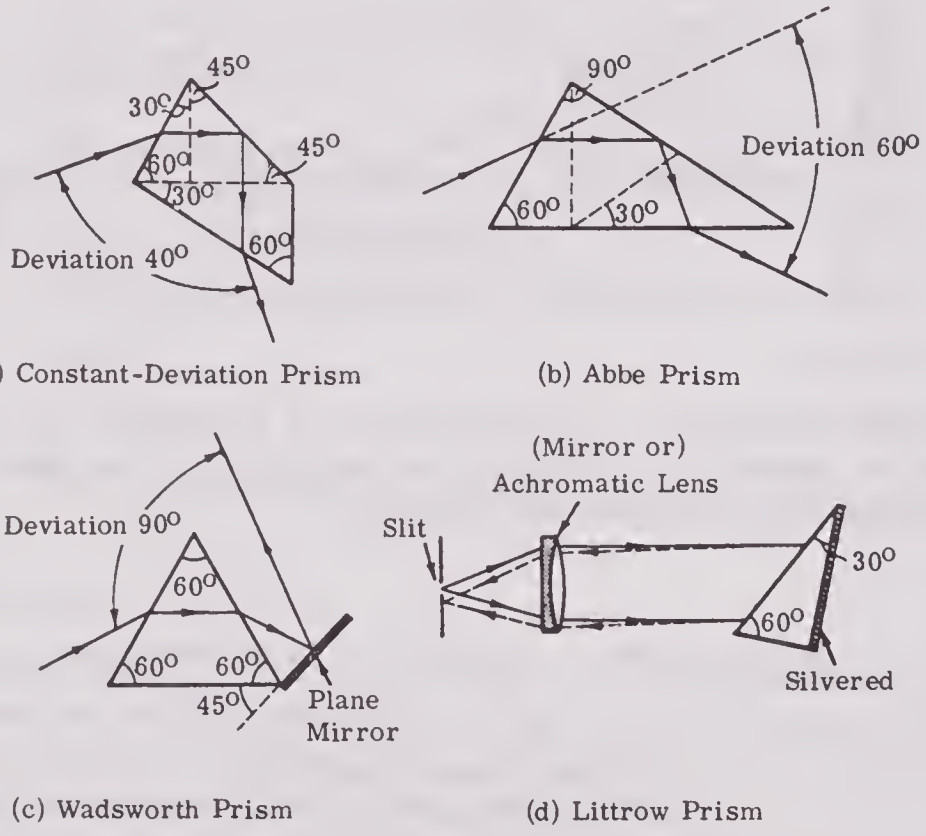


FIG. 7-31. Dispersing prisms.

7.11.2. Deviating Prisms [18]. Figure 7-32 shows various types of total reflecting deviating prisms. As shown in Fig. 7-32(*a*), rays enter perpendicular to one of the shorter faces of a total reflection prism, are totally reflected from the hypotenuse, and leave at right angles to the other short face. Such a prism can be used in two other ways, as shown in (*b*) and (*c*). The Dove prism, (*c*), interchanges the two rays, and if the prism is rotated about the direction of the radiation the rays rotate around each other with twice the angular velocity of the prism. The roof prism, (*d*), is similar

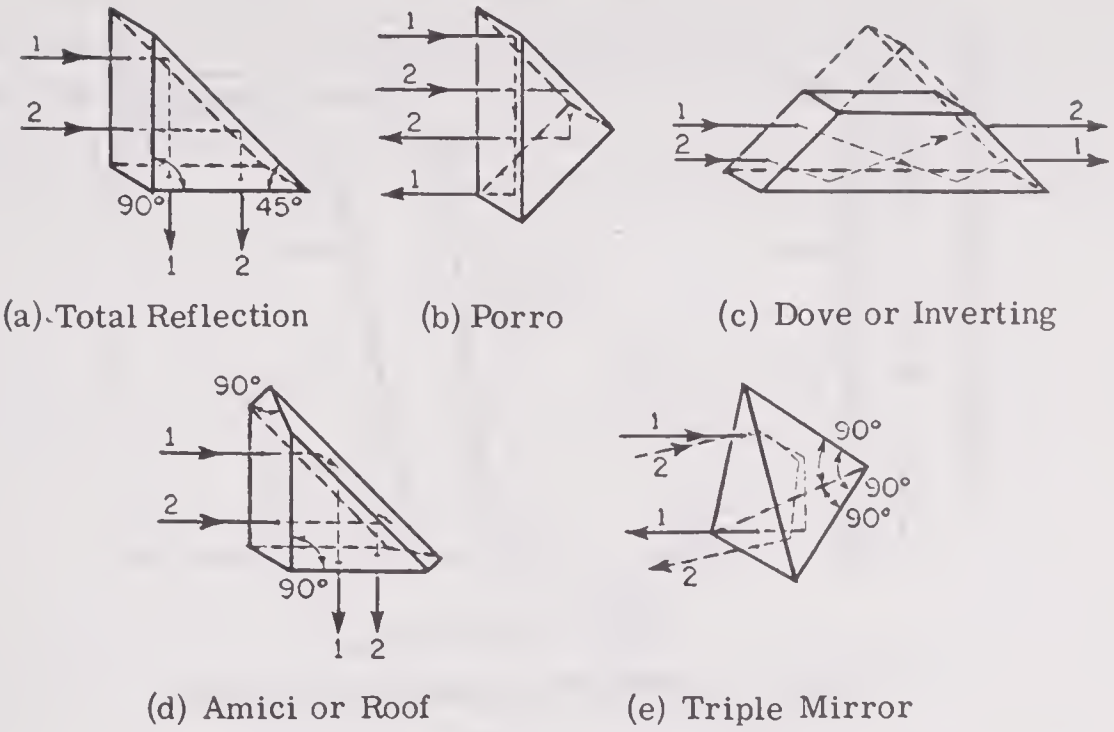


FIG. 7-32. Deviating prisms [18].

to the total reflection prism (*a*), except that it introduces an extra inversion. The triple mirror, (*e*), is made by cutting off the corner of a cube by a plane which makes equal angles with the three faces intersecting at that corner. It has the useful property that any ray striking it will, after being internally reflected at each of the three faces, be sent back parallel to its original direction.

7.11.3. Prism Materials. The most common prism materials for the infrared region include crystal quartz, rock salt, potassium bromide, lithium fluoride, and calcium fluoride. These materials are especially popular because of their availability as synthetic crystals (except for quartz, in general) [20].

Materials suitable for operation in the wavlength region beyond 15 μ, include silver chloride, thallium-bromo-iodide (KRS-5), cesium bromide, and cesium iodide.

The choice of an infrared prism material depends upon such characteristics as transmission, refractive index, and dispersion as a function of wavelength and possibly of temperature. In addition, the mechanical, physical, and chemical properties must also be considered. The properties of the aforementioned materials are discussed in Chapter 8.

7.12. Diffraction Gratings

A diffraction grating of the transmission type consists of a large number of small, equal-size, equally separated slits; each slit causes a diffraction pattern, and the waves from the individual slits also interfere to form a combined interference-diffraction pattern. The intensity can be written as

$$I = \frac{I_0}{\lambda^2} \frac{(\text{area})^2}{D^2} \frac{\sin^2\beta}{\beta} \frac{\sin^2 N\gamma}{\sin^2\gamma}$$

where $\beta = \dfrac{\pi a \sin\theta}{\lambda}$; a = slit width, θ = angle to image point

$$\gamma = \frac{\pi d \sin\theta}{\lambda}; \; d = \text{slit spacing}$$

I_0 = intensity at grating

D = grating-image distance

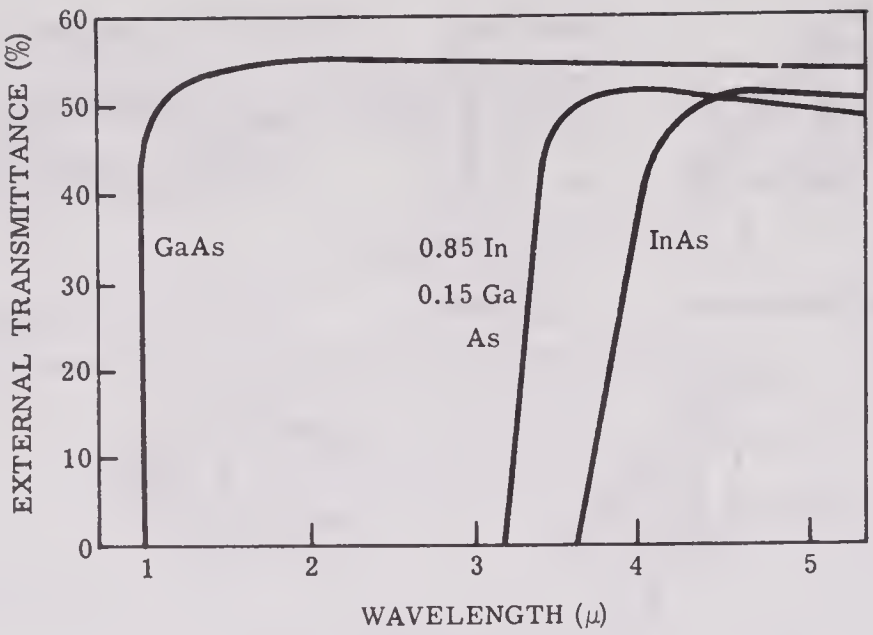


FIG. 7-33. Transmission of some mixed crystals.

The grating equation is

$$m\lambda = d(\sin\theta - \sin\theta')$$

where m = order of interference

θ = angle of incidence

θ' = angle of diffraction

d = slit separation

The angular dispersion is

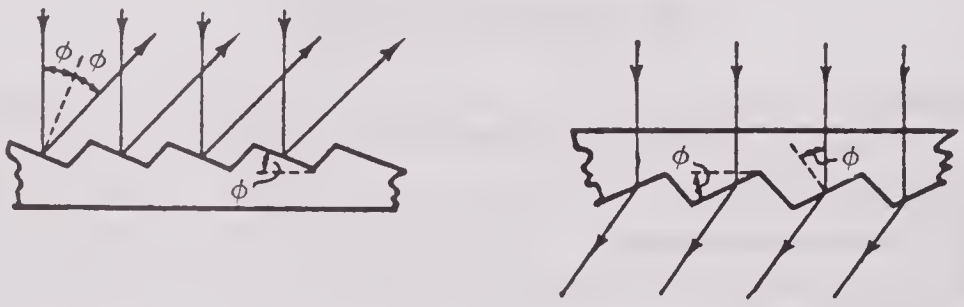
$$d\theta'/d = (m/d\cos\theta')$$

The resolving power is

$$\frac{\nu}{\lambda\nu} = (\lambda/d\lambda) = mN$$

where ν = wavenumber.

7.12.1. Blazed Gratings. It is possible to "blaze" a grating by ruling its grooves so that its sides reflect a large fraction of the incoming radiation of suitably short wavelengths in one general direction. Controlled groove shape is especially important in the gratings known as echelettes and echelles. In these gratings the grooves are ruled with one face optically flat. This face is inclined at an angle ϕ (see Fig. 7-34) to reflect or refract most of the incident radiation in a desired direction. In this way, the grating concentrates radiation in a particular spectral order, producing a brighter image than an ordinary diffraction grating.



(a) Reflecting Grating (b) Transmission Grating

FIG. 7-34. Reflection and transmission of gratings [18,19].

7.12.2. Concave Gratings. This type of grating consists of a concave mirror with ruled lines spaced equally along a chord across its center (Fig. 7-35). Radiation that passes through a slit and falls on such a grating is dispersed by it into spectra, in accordance with the standard grating formula. The concave grating requires no separate collimator or objective, since it is both a dispersing and a focusing element.

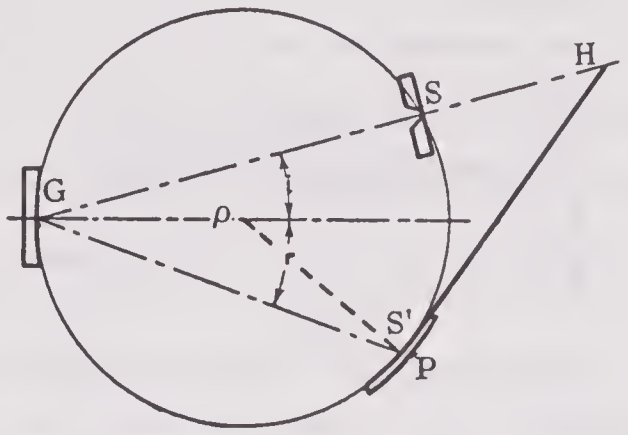


FIG. 7-35. Rowland circle focal curve of a concave grating [21].

Concave gratings have the advantage of focusing the radiation, thus eliminating the necessity for auxiliary lenses which introduce aberrations although the gratings suffer aberrations, principally astigmatism.

The focus for the spectrum from a concave grating is given by the formula

$$\left(\frac{\cos i}{\rho} - \frac{\cos^2 i}{S}\right) + \left(\frac{\cos r}{\rho} - \frac{\cos^2 r}{S'}\right) = 0$$

where i and r are the incident and reflected angles

ρ is the radius of curvature of the concave blank

S is the slit-to-grating distance

S' is the distance from the grating to the image of the spectrum

This equation can be satisfied by setting $S = \rho \cos i$ and $S' = \rho \cos r$, the polar equations for a circle of diameter ρ, containing the points S, S', and G, which is shown in Fig. 7-35 as the familiar Rowland circle.

7.12.3. Ebert-Fastie Plane Grating Mountings [21]. Figure 7-36(a) shows the original Ebert [22] mounting, which was modified by Fastie [23] as shown in Fig. 7-36(b). Both mountings can be designated as "side-by-side" designs; the entrant rays are on one side of the grating, the emergent rays on the other. Figure 7-36(c) is an "under-over" design. The entrant rays pass below the grating and the emergent rays pass above.

7.12.4. Concave Grating Mountings [21]. Figure 7-37 shows different mountings for the concave grating designed to maintain slit, grating, and plate- or film-holder on the Rowland circle (Fig. 7-36).

The Rowland mounting, Fig. 7-37(a), utilizes the geometric principle that the locus of the apex of a series of right triangles having a common hypotenuse is a circle having the hypotenuse as a diameter. The grating, G, and plateholder, P, are fixed on opposite ends of a rigid bar, which forms the hypotenuse and the diameter of the Rowland circle. The slit, S, is placed at the intersection of two tracks, SP and SG, which are carefully constructed at right angles. Rollers under G and P are constrained to travel along SG and SP, so that as the bar, GP, is moved, the grating remains on the optical

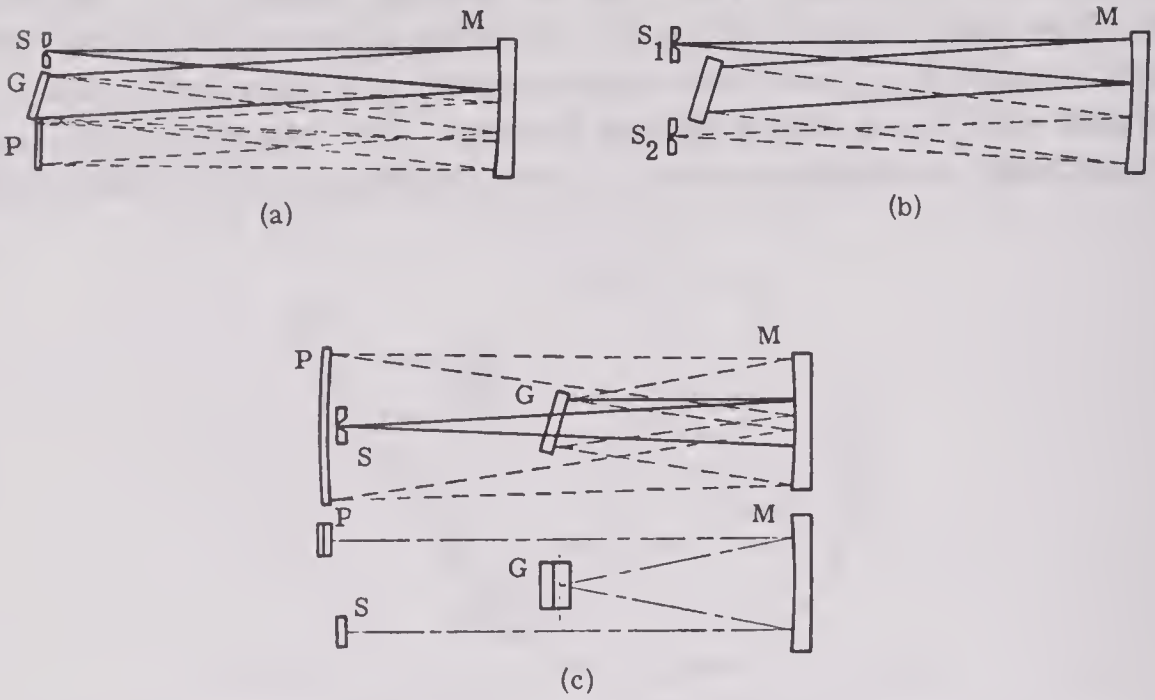


FIG. 7-36. Plane grating mountings; (*a*) original Ebert design; (*b*) Fastie modified "side-by-side" design; (*c*) Fastie modified "over-under" design [21].

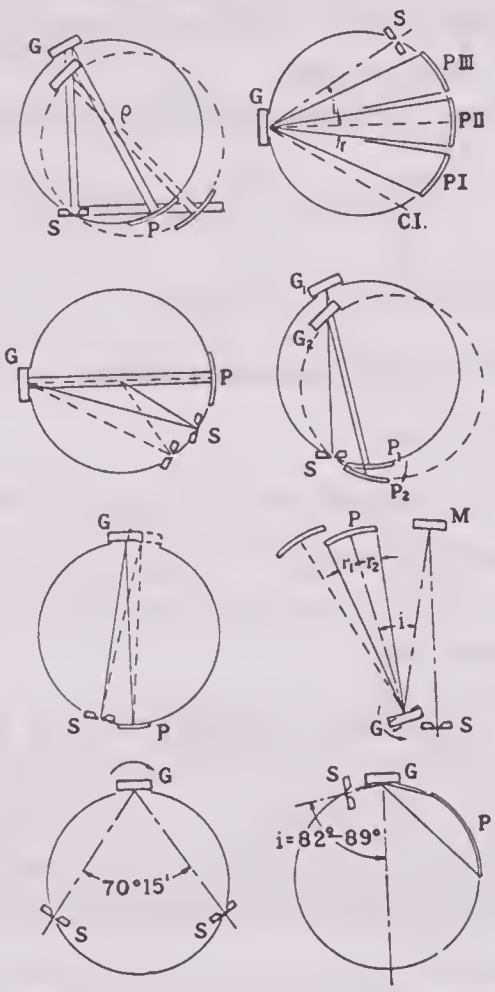


FIG. 7-37. Concave grating mountings [21].

axis fixed by the condensing system, and the Rowland circle moves to coincide with all three optical elements [24].

The Paschen-Runge or Paschen mounting, Fig. 7-37(*b*), usually consists of large circular tracks 21 or 35 feet in diameter, which are built in a room that can be temperature stabilized and darkened. Plateholders can be clamped to these tracks where desired. The slit mounts protrude through a wall into an adjoining room where the excitation equipment and light sources are located [25].

In the Abney mounting, Fig. 7-37(*c*), the plateholder is mounted on the normal to the grating. The slit is mounted on an arm rotating about the center of the Rowland circle to change the wavelength range. The slit assembly must also be rotated about an axis below the slit opening itself, so that the source-slit axis will still point at the grating [26].

For the Eagle mounting, Fig. 7-37(*d*), three independent mechanical adjustments are required to maintain the elements on the Rowland circle.

(1) The grating must be rotated to change the wavelength range.

(2) The grating must be moved along the optical axis to maintain the focus.

(3) The plateholder must be tilted about an axis under the slit to remain in focus across its length [27].

In the Beutler radius mounting, Fig. 7-37(*e*), the slit and plateholder are mounted permanently on the Rowland circle, and the grating is mounted on an arm which rotates about the center of the Rowland circle [28].

For the Wadsworth mounting, Fig. 7-37(*f*), the focal distance is half the radius of curvature, the individual elements are not located on the Rowland circle, and a second optical element is introduced in the form of a concave collimating mirror. This arrangement produces a stigmatic image and a linear dispersion, as in the Rowland and Abney mountings. The concave mirror is mounted at its focal distance from the slit, tilted slightly off axis to irradiate the grating with parallel light. The grating may be located either side of the optical axis, as close to it as possible to minimize aberrations. The plateholder is located on a bar whose axis forms the normal to the grating. To change the range to higher wavelengths, the bar is rotated away from the mirror about an axis under the grating face. The plateholder must be moved away from the grating to remain in focus. In any position of the plateholder the shorter wavelength radiation is nearest to the concave mirror [29].

Seya-Namioka mounting, Fig. 7-37(*g*), is based on the principle that if the angle between the entrant and emergent rays in a spectrometer is 70° 15′ there will be such slight defocusing, if the grating is merely rotated about its own vertical axis, that the image would be entirely acceptable for scanning monochromator usage [30].

Grazing incidence mounting, Fig. 7-37(*h*), [31], operates on the principle that, if a grating is illuminated at grazing incidence, the short-wavelength (below 1000 A) radiation will be totally reflected. Angles of incidence as high as 85-89° have been employed to observe wavelengths as low as 53 A [32] or even 12.1 A [33].

7.12.5. Production of Gratings. Gratings are engraved by highly precise ruling engines which use a diamond tool to press a series of many thousands of fine shallow burnished grooves into a smooth metallic surface.

Gratings for the range 1500 to 10,000 A are commonly ruled with 5000 to 30,000 grooves per inch (the usual value is near 15,000), on a thin layer of aluminum deposited on glass by evaporation in vacuum. Gratings for the infrared region are also ruled on gold, silver, copper, lead, or tin mirrors with coarser groove spacings.

Gratings of 2-in., 4-in., or 6-in. ruled width are commonly used in commercial spectrographs, with projection distances of 20-180 in. In large research instruments, gratings of 6- to 10-in. ruled width are used with projection distances of 10-50 ft or more. The largest modern gratings, used in their highest orders, show resolving power $\lambda/\delta\lambda$ in excess of 900,000 in the green region of the spectrum, and in excess of 1.5×10^6 at shorter wavelength. Here λ is the mean wavelength of two closely spaced, just-resolvable spectral lines and $\delta\lambda$ is their wavelength difference.

References

1. "Data Sheets," Eastman Kodak Co., Rochester, N. Y. (1962).
2. "Data Sheets," Servo Corp. of America, Hicksville, N. Y. (1962).
3. "Advanced Optics Engineering, Design and Fabrication," Bulletin 6-010, Barnes Engineering Corp., Stamford, Conn.
4. "Near-Infrared Transmission Filters," Progress Report No. 3, Bausch & Lomb Optical Co., Rochester, N. Y. (1958).
5. C. F. Mooney and A. F. Turner, "Infrared Transmitting Interference Filters, Proceedings of the Conference on Infrared Optical Materials, Filters, and Films," Engineer Research and Development Laboratories, Fort Belvoir, Va. (1955).
6. G. Hass and A. F. Turner, *Coatings for Infrared Optics,* Reprint from Wissenschaftliche Verlagsgesellschaft m.b.h. Stuttgart, 143-163.
7. R. A. Smith, F. E. Jones, and R. P. Chasmar, *The Detection and Measurement of Infrared Radiation,* Oxford University Press (1957).
8. *Engineering Notebook,* Astrionics Division, Aerojet-General Corp., Azusa, Calif. (1961).
9. E. D. McAlister, "The Christiansen Light Filter; Its Advantages and Limitations," Smithsonian Miscellaneous Collections, **93**, 7 (1936).
10. R. B. Barnes and T. W. Bonner, *Phys. Rev.,* **49**, 732 (1936).
11. J. Strong, *Concepts of Classical Optics,* W. H. Freeman and Co., San Francisco (1958).
12. "Data Sheets," Infrared Industries, Waltham, Mass. (1962).
13. N. M. Mohler and J. R. Loofbourow, "Optical Filters," *Am. J. Phys.,* **20**, 579-588 (1952).
14. J. H. Shenk *et al., J. Opt. Soc. Am.,* **36**, 10, 569 (1946).
15. W. L. Wolfe and S. S. Ballard, "Optical Materials, Films, and Filters for Infrared Instrumentation," *Proc. I.R.E.,* **47**, 9 (1959).
16. R. H. Bube, *Photoconductivity of Solids,* Wiley, N. Y. (1960).
17. M. R. Holter, S. Nudelman, G. H. Suits, W. L. Wolfe, and G. J. Zissis, *Fundamentals of Infrared Technology,* Macmillan, N. Y. (1962).
18. F. A. Jenkins and H. E. White, *Fundamentals of Optics,* 2nd ed. McGraw-Hill, N. Y. (1950).
19. H. L. Hackforth, *Infrared Radiation,* McGraw-Hill, N. Y. (1960).
20. V. Z. Williams, "*Infrared Instrumentation Techniques,*" *Rev. Sci. Instr.,* **19**, 135-178 (1948).
21. G. L. Clark, *The Encyclopedia of Spectroscopy,* Reinhold, N. Y. (1960).
22. H. Ebert, *Wied. Ann.,* **38**, 489 (1889).
23. W. G. Fastie, *J. Opt. Soc. Am.,* **42**, 9, 641 (1952).
24. H. A. Rowland, *Phil. Mag.,* **16**, 197 and 210 (1883).
25. C. R. Runge and F. Paschen, *Abhandl. Akad. Wiss. Berlin,* Anhang 1 (1902).
26. W. DeW. Abney, *Phil. Trans.,* **177**, II, 457 (1886).
27. A. Eagle, *Astrophys. J.,* **31**, 120 (1910).
28. R. A. Sawyer, *Experimental Spectroscopy,* Prentice-Hall, N. Y. (1944).
29. F. L. O. Wadsworth, *Astrophys. J.,* **3**, 54 (1896).
30. M. Leya, *Sci. of Light,* **2**, 8, Tokyo (1952).
31. J. B. Hoag, *Astrophys. J.,* **66**, 225 (1927).
32. B. Edlen, *Nova Acta Regiae Soc. Sci. Upsaliensis,* **9**, 6 (1934).
33. F. Tyren, *Z. Physik.,* **III**, 314 (1938).

Chapter 8

OPTICAL MATERIALS

William L. Wolfe

The University of Michigan

CONTENTS

8.1. Introduction 317
8.2. Types of Materials 317
8.2.1. Glasses 317
8.2.2. Semiconductor Crystals 324
8.2.3. Plastics 325
8.2.4. Some Recent Hot-Pressed Samples 326
8.3. Comparisons of Material Properties 326
8.3.1. Transmission Region 326
8.3.2. Refractive Index 328
8.3.3. Dispersion 328
8.3.4. Dielectric Constant 329
8.3.5. Melting Temperature 330
8.3.6. Thermal Conductivity 331
8.3.7. Thermal Expansion 331
8.3.8. Specific Heat 333
8.3.9. Hardness 333
8.3.10. Solubility 333
8.3.11. Young's Modulus 333
8.4. Material Data Useful for Lens Design 335
8.4.1. Dispersion Equations for Individual Materials 335
8.4.2. Herzberger Dispersion Equation 337
8.4.3. Refractive Index, Dispersion Values 338
8.4.4. G Sums 341
8.5. Useful Equations 347
8.5.1. Reflection and Transmission at a Single Surface 347
8.5.2. Reflection and Transmission by a Single Layer 348
8.5.3. Emissivity, Transmissivity, and Reflectivity for Partially Transparent Bodies 349
8.5.4. Loss Tangent 350
8.5.5. Extinction Coefficient 351

8.6. Optical Surface Coatings 351
8.6.1. Reflective Coatings 351
8.6.2. Filter Mirrors 352
8.6.3. Protective Coatings 354
8.6.4. Antireflection Coatings 355
8.6.5. Replica Mirrors 355
8.7. Radiation Damage 358
8.8. Infrared Transmission of Cooled Optical Materials 359
8.9. Optical Properties of Blacks 359
8.10. Optical Properties of Water 362
8.11. Effects of Space Radiation on Optical Materials 364

8. Optical Materials

8.1. Introduction

Data on optical, mechanical, thermal, and chemical properties of infrared optical materials are presented here. For additional information see [1], publications of the various manufacturers of optical materials, and the references.

8.2. Types of Materials

Optical materials in current use include glasses, crystals, plastics, and hot-pressed materials. The crystals, natural and artificial, are dielectric and semiconducting.

8.2.1. Glasses. Most optical glasses transmit into the infrared only to a wavelength of about 2.7 μ and then fairly strong absorption sets in; beyond about 4 μ, glasses have no appreciable transmission. Figure 8-1 shows the transmission of several infrared-transmitting glasses that do not demonstrate a strong "water-band" absorption at 2.8 μ but exhibit a slower dropoff in transmission in the 3-μ to 4-μ region. Figure 8-2 is the standard n-ν curve for ordinary glasses, that is, the refraction for the D-line plotted against $\nu = n_{D-1}/(n_F - n_C)$.

8.2.1.1. *High-Silica Glasses and Fused Quartz.* The names fused quartz, silica glass, fused silica, etc., now mean about the same thing: amorphous mixtures that are almost entirely silica. Originally the names differentiated among natural and synthetic materials. Some special silicate glasses recover transmission beyond the 2.8-μ

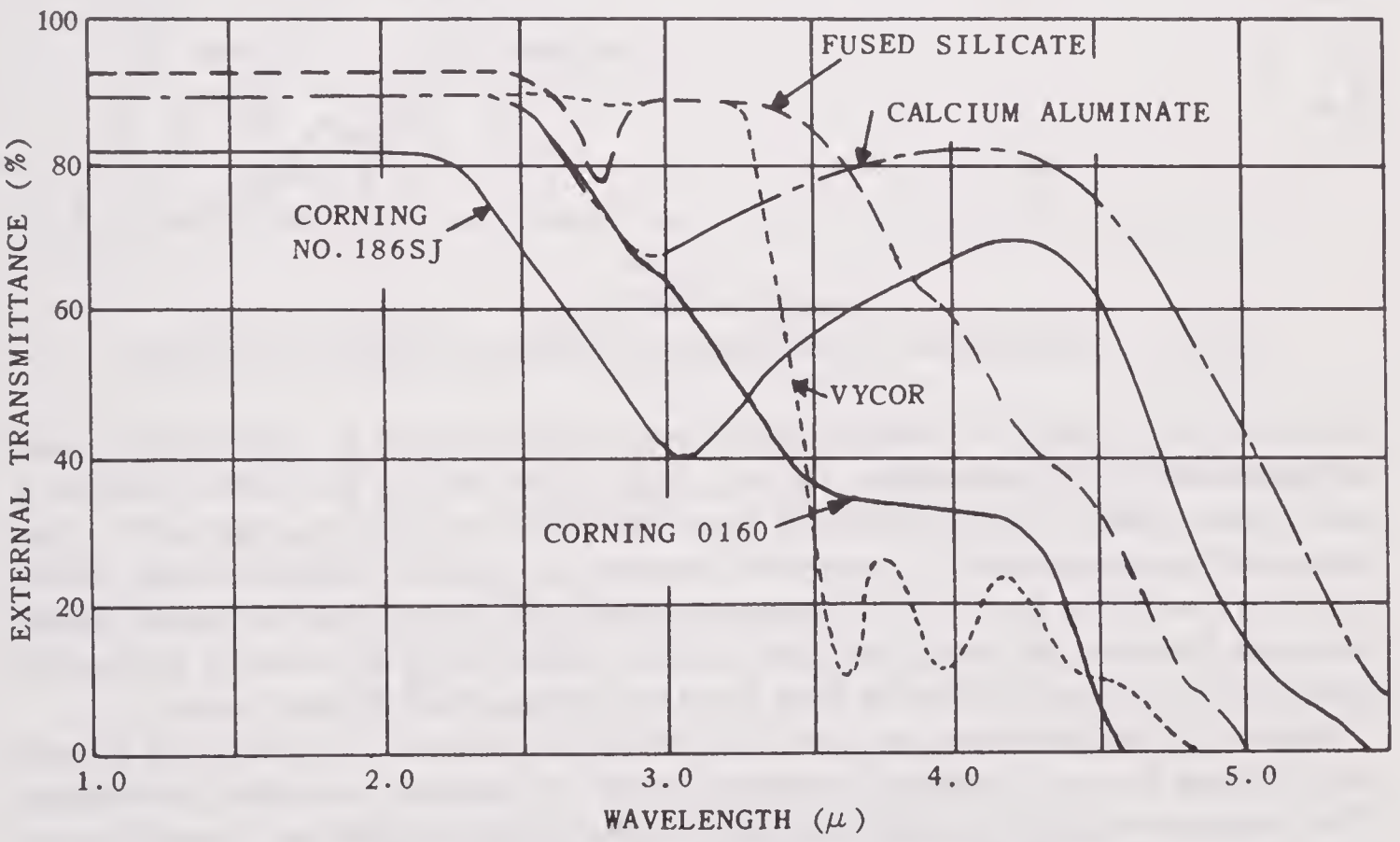


FIG. 8-1. The transmission of several infrared transmitting glasses; thickness, 2 mm.

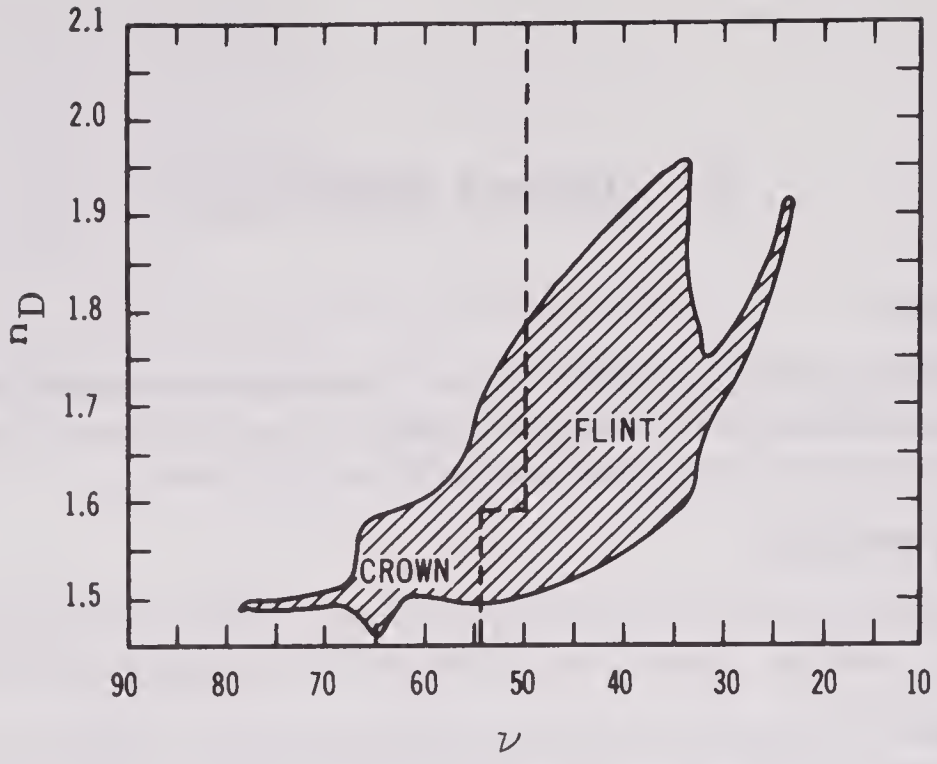


FIG. 8-2. Dispersion region for standard glasses.

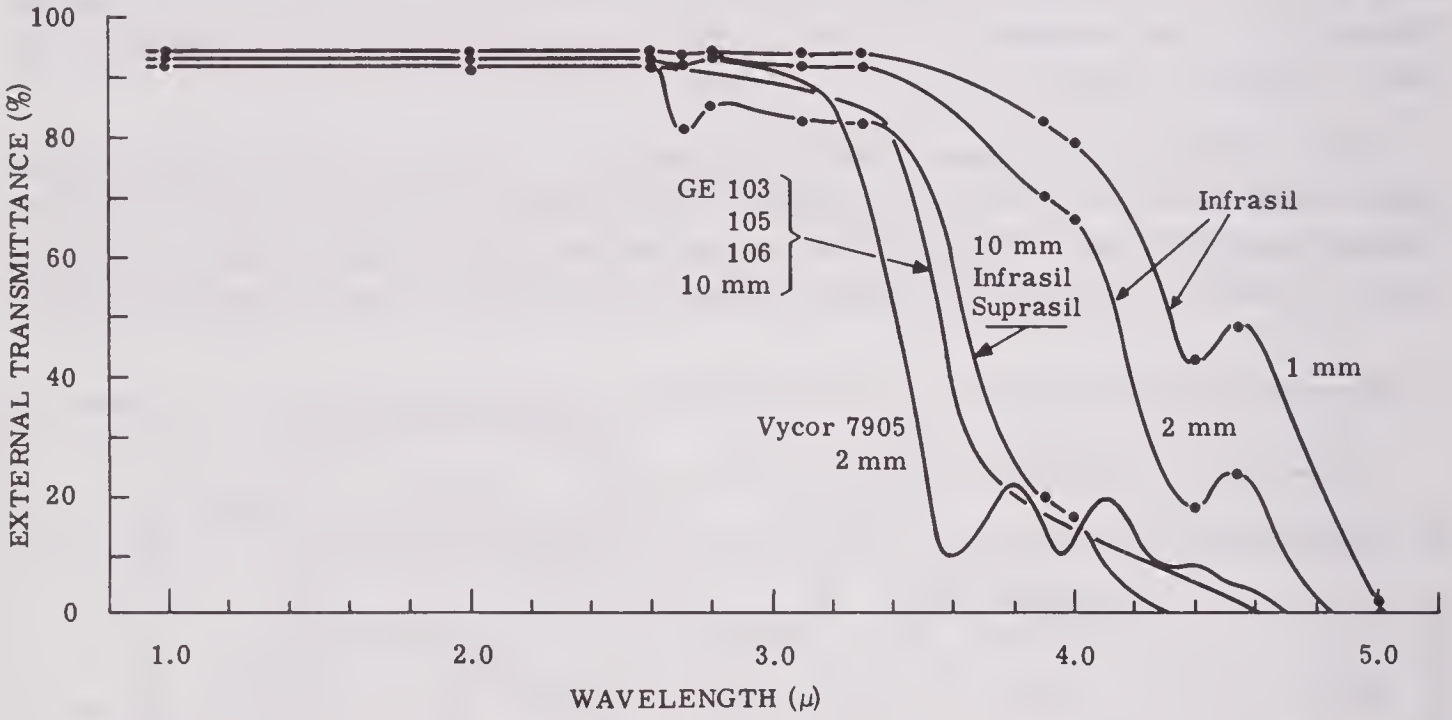


FIG. 8-3. The transmission of some high-silica glasses (from company data sheets).

absorption band and give reduced transmission out to about 5 μ. This performance is illustrated by the transmission curve in Fig. 8-1 for Corning No. 186SJ, which is a lead silicate glass. The properties of fused and crystal quartz are given in [1]. Examples of the properties of a number of commercial varieties (obtained from manufacturers' brochures and private communications) are given below to provide useful data and illustrate the small variation in most of the physical constants among the glasses and the drastic difference from these in the properties of fused quartz.

Most of the useful physical properties are listed in Table 8-1. Variations in glasses are obtained by use of different purifications and by different annealing procedures. The transmission of the samples from the principal suppliers (Corning, Dynasil, Engelhard, General Electric, Heraeus, and Sawyer) are given in Fig. 8-3. The 2-mm Vycor sample is from Corning; Infrasil is from Engelhard; GE 103, 105, and 106 are from the General Electric Lamp Glass Department (the GE 102 and 104 have in addition to

Table 8-1. Properties of Silica Glasses

Material	*Density* (g/cc)	*Softening Temp.* (°C)	*Strain Temp.* (° C)	*Specific Heat*	*Thermal Conductivity* [10^{-3} cal × (cm sec C°)$^{-1}$]	*Thermal Expansion* [10^{-6} (C°)$^{-1}$]	*Young's Modulus* (10^6 psi)	*Modulus of Rupture* (psi)	*Modulus of Rigidity* (psi)	*Knoop No.*	*Dielectric Constant* (1 Mc)
Vycor 7905	2.18	1500	820	0.255	3.3	0.8	9.6	7.3	–	477	3.8
GE 103	2.20	1667	1079	–	3.3	0.55	10.4	–	4.49	–	–
GE 101, 102, 104, 105, 106	2.20	1667	1079	–	3.3	0.55	–	–	–	–	–
Engelhard Infrasil	2.2	1700	–	–	3.3	0.55	10.4	–	–	470	–
Crystal Quartz	2.65	1470 α-β 575°	–	0.188	20.1	10.67	12.6	–	–	741	4.3 at 10 Mc

TABLE 8-2. PROPERTIES OF SILICATE, GERMANATE, AND ALUMINATE GLASSES

Material	*Density* (g/cc)	*Softening Temp.* (°C)	*Strain Temp.* (°C)	*Specific Heat*	*Thermal Conductivity* [10^{-3} cal × (cm sec C°)$^{-1}$]	*Thermal Expansion* [10^{-6} (C°)$^{-1}$]	*Young's Modulus* (10^6 psi)	*Modulus of Rupture* (psi)	*Modulus of Rigidity* (psi)	*Knoop No.*	*Dielectric Constant* (1 Mc)	*Loss Tangent* (1 Mc)
Germanates												
Corning												
9752	4.51	785	629	0.14	–	8.3	8.74	5520	–	473	13	0.001
0160	3.05	610	388	0.17	–	9.7	9.1	–	–	460	7	0.001
Silicates												
NBS F158	4	836	–	–	–	8.4	15.5	–	–	–	–	–
Aluminates												
Bausch & Lomb												
RIR-10, -11, -12	3.07	900	–	–	–	8.1	15.2	–	–	594	9	0.005
RIR-2	4.57	800	–	–	3.2	9.6	10.2	–	–	576	9	0.005
RIR-20	5.18	760	–	–	–	8.4	12-14	–	–	542	9	0.005
GE												
Lucalox	3.98	2040	–	–	0.1	7	56.1	–	23.3	–	9.9	–
Barr & Stroud												
BS 37B	–	–	–	–	–	–	–	–	–	–	–	–

these curves a narrow, deep 2.8-μ absorption); Suprasil is the best of the Heraeus fused silicas; the Sawyer material is virtually identical to Infrasil; and the Dynasil is similar to Infrasil but with a deep absorption at 2.8 μ. Sometimes this absorption can be alleviated by careful attention to water elimination during preparation.

8.2.1.2. *Silicate Glasses.* The National Bureau of Standards in its studies of infrared-transmitting glasses has produced some silicates with fairly good transmission beyond the OH band at 2.8 μ. One example is given in Table 8-2. Actually this is a composite set of properties (taken from the NBS report series) and indicative of no particular glass. The properties of individual samples will *approximate* these values but change with variation in composition.

8.2.1.3. *Germanate Glasses.* The properties of several varieties of germanate glasses are displayed in Tables 8-2 and 8-3 and Figs. 8-4 and 8-5. The properties of individual samples will approximate the values given but change with variation in composition.

8.2.1.4. *Calcium Aluminate Glasses.* Bausch and Lomb RIR-2, -10, -11, -12, and -20, GE Lucalox, and Barr and Stroud 36A (old type) and 37B are all types of calcium aluminate glass. Properties are given in Tables 8-2 and 8-4 and Fig. 8-6. The 2.8-μ absorption band can be eliminated by careful attention to humidity during preparation, *e.g.*, by vacuum melting.

8.2.1.5. *Nonoxide Glasses.* Arsenic-modified selenium glass, arsenic-sulfur glass, and other nonoxide glasses are not restricted to those containing the silicate or aluminate radical. They are described in [1]. Some newer materials include mixtures of Se, Te, S, As, Ge, and a few other heavy atoms. Reports on these materials can be found in [2] and [3].

TABLE 8-3. REFRACTIVE INDEX OF NATIONAL BUREAU OF STANDARDS F998

Wavelength (μ)	*Germanate Glasses F998*	*Wavelength* (μ)	*Germanate Glasses F998*
0.4358	1.88997	1.9701	–
0.4861	1.87470	2.1526	1.81364
0.5461	1.86242	2.2493	–
0.5780	1.85755	2.3126	1.81151
0.5893	1.85597	2.4374	1.81000
0.6439	1.84986	2.5947	1.80797
0.6563	1.84866	2.6685	1.80680
0.8521	1.83632	2.998	–
1.0140	1.83082	3.3033	1.79706
1.1287	1.82801	3.422	1.79490
1.3622	1.82374	3.5078	1.79345
1.5295	1.82142	4.225	–
1.6606	–	4.253	1.77863
1.6932	1.81930	4.281	–
1.7012	–	5.138	1.75437
		5.343	1.74805

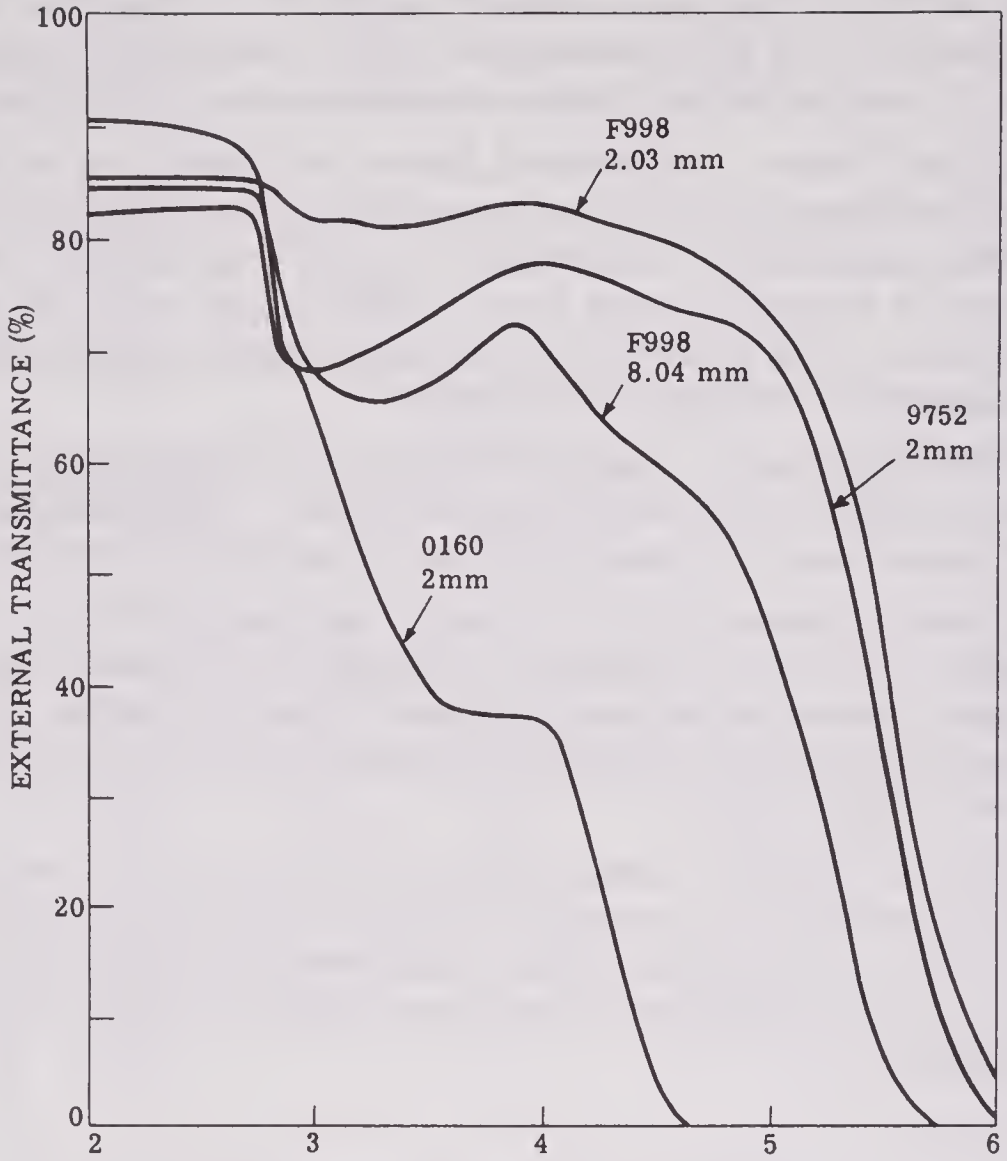


FIG. 8-4. Transmission of several germanate glasses (from company data sheets and NBS reports).

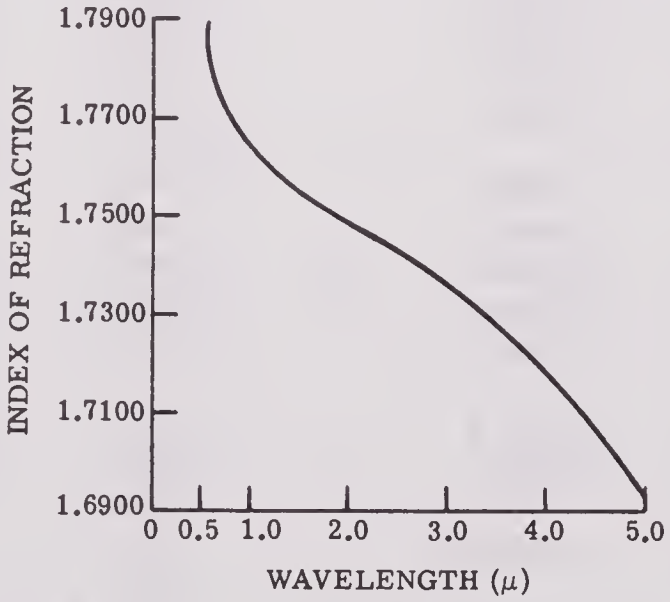


FIG. 8-5. Index of refraction of Corning No. 9752 glass.

TABLE 8-4. REFRACTIVE INDEX OF BAUSCH AND LOMB CALCIUM ALUMINATE GLASSES

Wavelength (μ)	RIR-10	RIR-11	RIR-12	RIR-2	RIR-20
0.4047	–	–	–	1.82800	1.91449
0.4341	–	–	–	1.81806	1.90155
0.4359	–	–	–	1.81746	1.90082
0.4861	1.66057	1.67887	1.66647	1.80536	1.88529
0.5461	1.65385	1.67109	1.65919	1.79558	1.87274
0.5876	–	–	–	1.79060	1.86639
0.5893	1.65022	1.66699	1.65532	1.79041	1.86616
0.6563	1.64588	1.66239	1.65085	1.78443	1.85866
1.0140	1.6352	1.6506	1.6397	1.76988	1.84044
1.1287	1.6334	1.6486	1.6378	1.76741	1.83762
1.3620	1.6304	1.6455	1.6346	1.76343	1.83333
1.5295	1.6285	1.6435	1.6328	1.76104	1.83000
1.6606	1.6271	1.6420	1.6313	1.75920	1.82909
1.8131	1.6255	1.6404	1.6297	1.75718	1.82722
1.9701	1.6238	1.6386	1.6280	1.75503	1.82527
2.1526	1.6216	1.6364	1.6259	1.75238	1.82290
2.2493	–	–	–	1.75103	1.82183
2.3254	1.6196	1.6344	1.6239	1.74984	1.82073
2.4374	1.6182	1.6329	1.6224	1.74806	1.81924
2.577	–	1.6310	1.6206	1.74582	1.81732

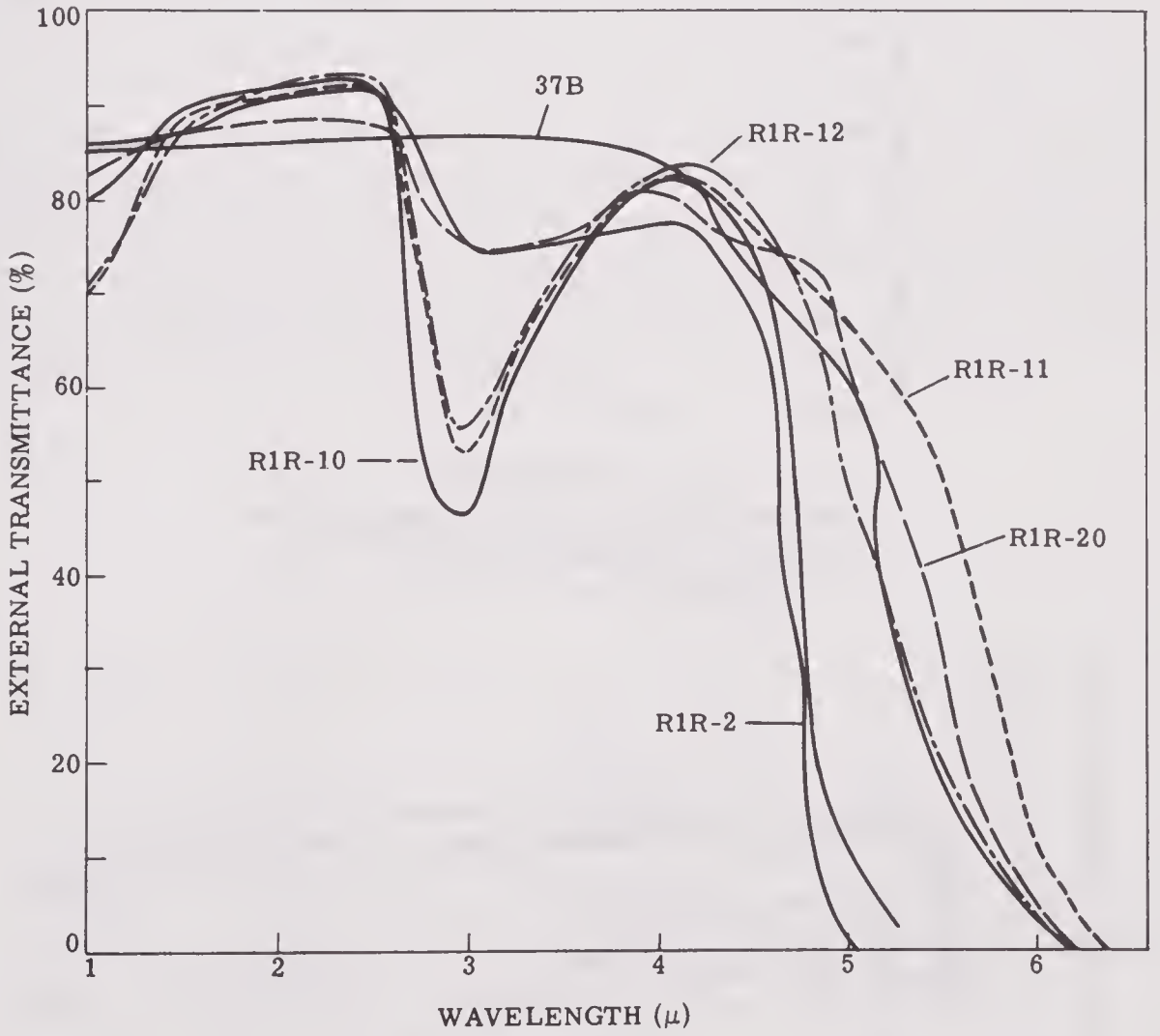


FIG. 8-6. Transmission of several calcium aluminate glasses; thickness, 2 mm (from company data sheets and NBS reports).

8.2.2. **Semiconductor Crystals.** The cuton wavelength and the transmittance of semiconductors are functions of temperature (Fig. 8-7) and purity (Fig. 8-8). Figure 8-9 shows the transmission of various semiconducting materials.

By their basic nature semiconducting materials have small energy gaps which correspond to cuton wavelengths in the infrared. Increasing the temperature of the material in effect narrows the gap, thereby increasing the cuton wavelength. It also increases the probability that electrons can have energies characteristic of the conduction band. Many additional data will be found in [1] and in the several journals on solid-state physics.

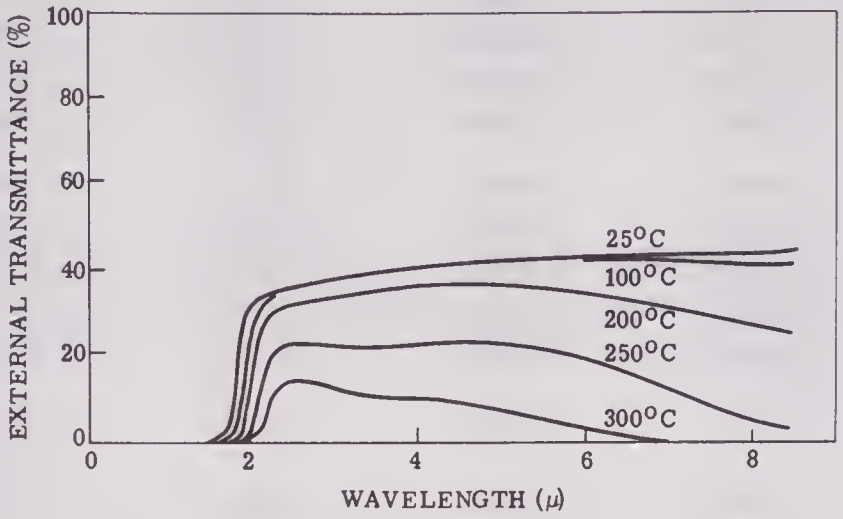


FIG. 8-7. Effect of temperature on semiconductor transmission; sample is 1.17 mm, 30 Ω cm Ge.

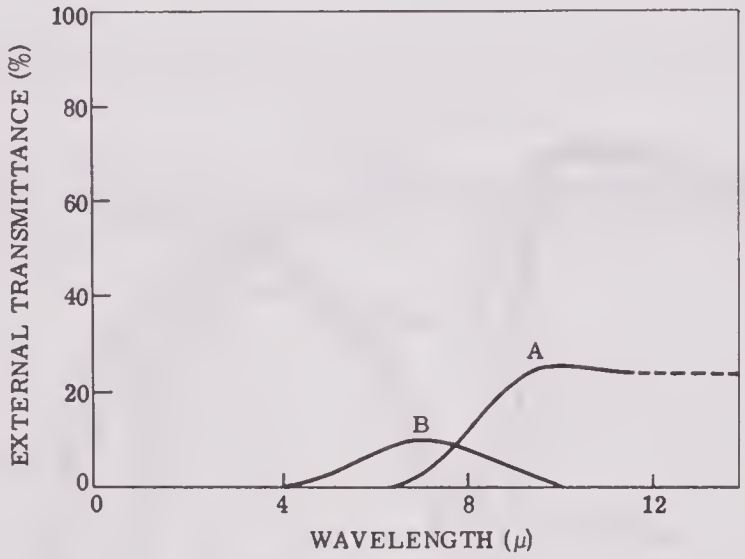


FIG. 8-8. Effect of purity on semiconductor transmission; samples are 0.2 mm, 7×10^{-3} Ω cm InSb and 0.2 mm, 2.5×10^{-4} Ω cm InSb.

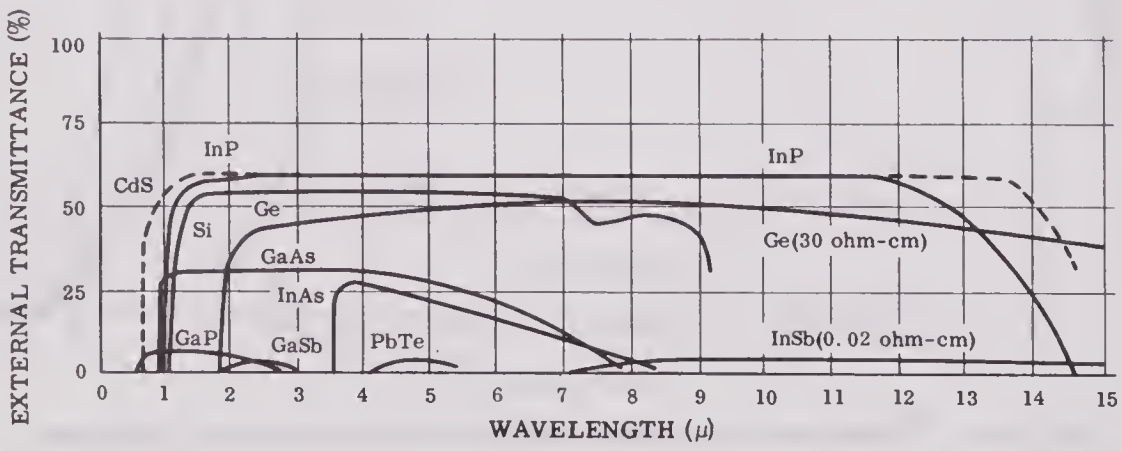


FIG. 8-9. Transmission of selected semiconductor materials.

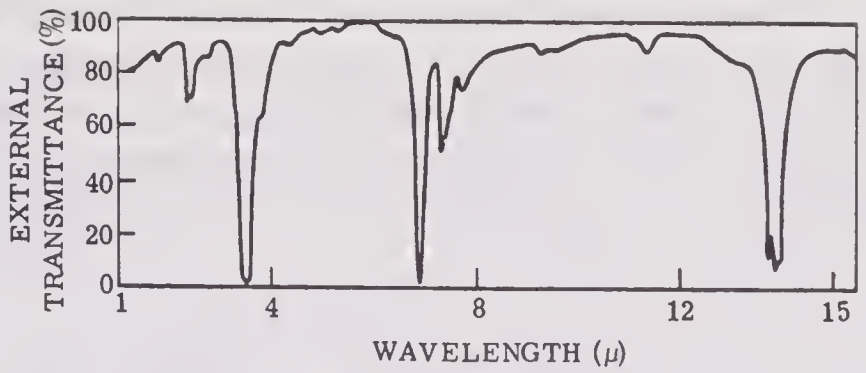


FIG. 8-10. Transmission of polyethylene; thickness, 0.1 mm.

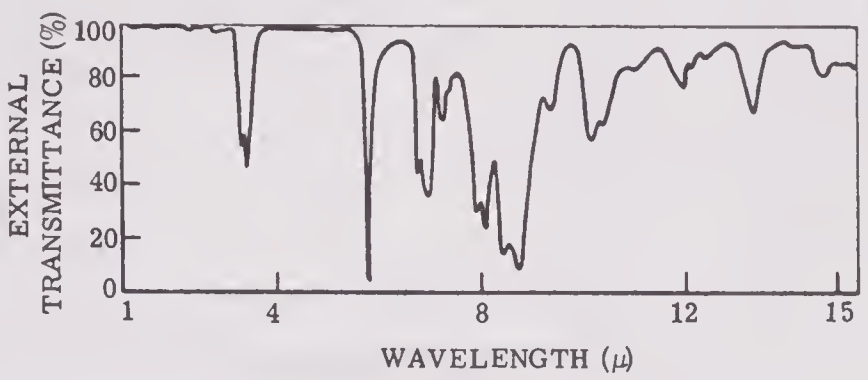


FIG. 8-11. Transmission of Plexiglas; thickness, 0.2 mm.

8.2.3. Plastics. Among the common plastics used for infrared applications are polyethylene and polymethylmethacrylate (available commercially as Lucite or Plexiglas respectively). Transmission curves of these two materials are shown in Figs. 8-10 and 8-11. These curves are representative of the transmission of thin films of many different plastics. Except for narrow bands, where only a small amount of energy is absorbed, and for broader bands in some materials, the transmission is relatively good. For thicker samples, however, the regions of small absorption deepen rapidly and widen considerably, and the absorption becomes so great that the material may no longer be satisfactory for the intended use. Table 8-5 shows the wavelengths at which various molecules and molecular groups absorb infrared radiation. These wavelengths are the characteristic absorption bands of the bonds or groups noted.

TABLE 8-5. ABSORPTION WAVELENGTHS OF CHARACTERISTIC GROUPS [1]

Molecular Group	*Wavelengths of Absorption* (μ)
O – H	2.8
N – H	3.03, 6.12, 6.46
C – H	3.4, 6.81
Carbonyl	5.75
Methyl	7.26, 7.71
Ester	8, 9.74
C – O	9(broad), 13.47-14.20
C – Cl	14-15
Si – O	4.5, 9.5
Al – O	~6.5
Ge – O	~5.5, 11.6

Kel-F, a polymer of trifluorochloroethylene, is used for windows and coatings although in thicknesses greater than about 0.25 in. it is difficult to manufacture with suitably high transmission. It has a low thermal conductivity, and its refractive index may vary throughout a sample. The transmittances of two different thicknesses of Kel-F are shown in Fig. 8-12.

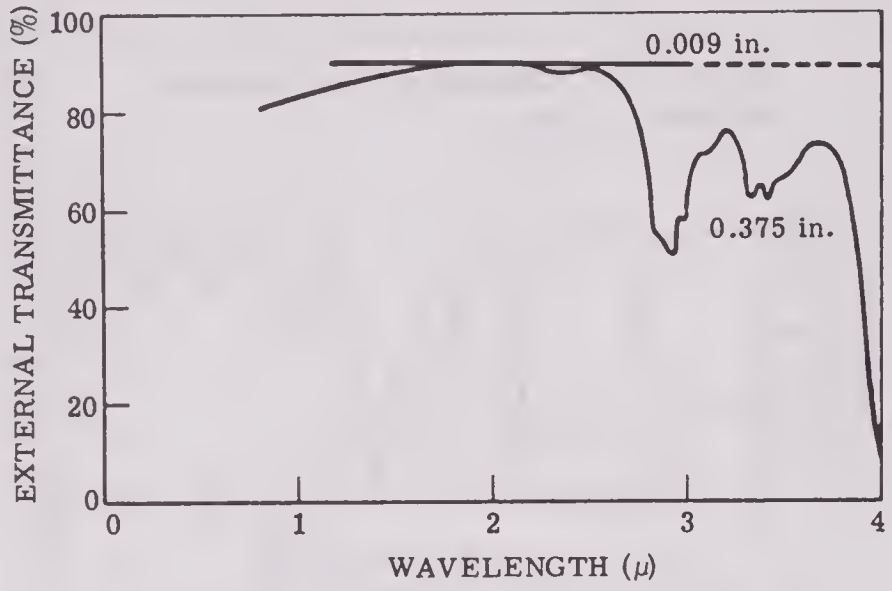


FIG. 8-12. Transmission of Kel-F for two thicknesses.

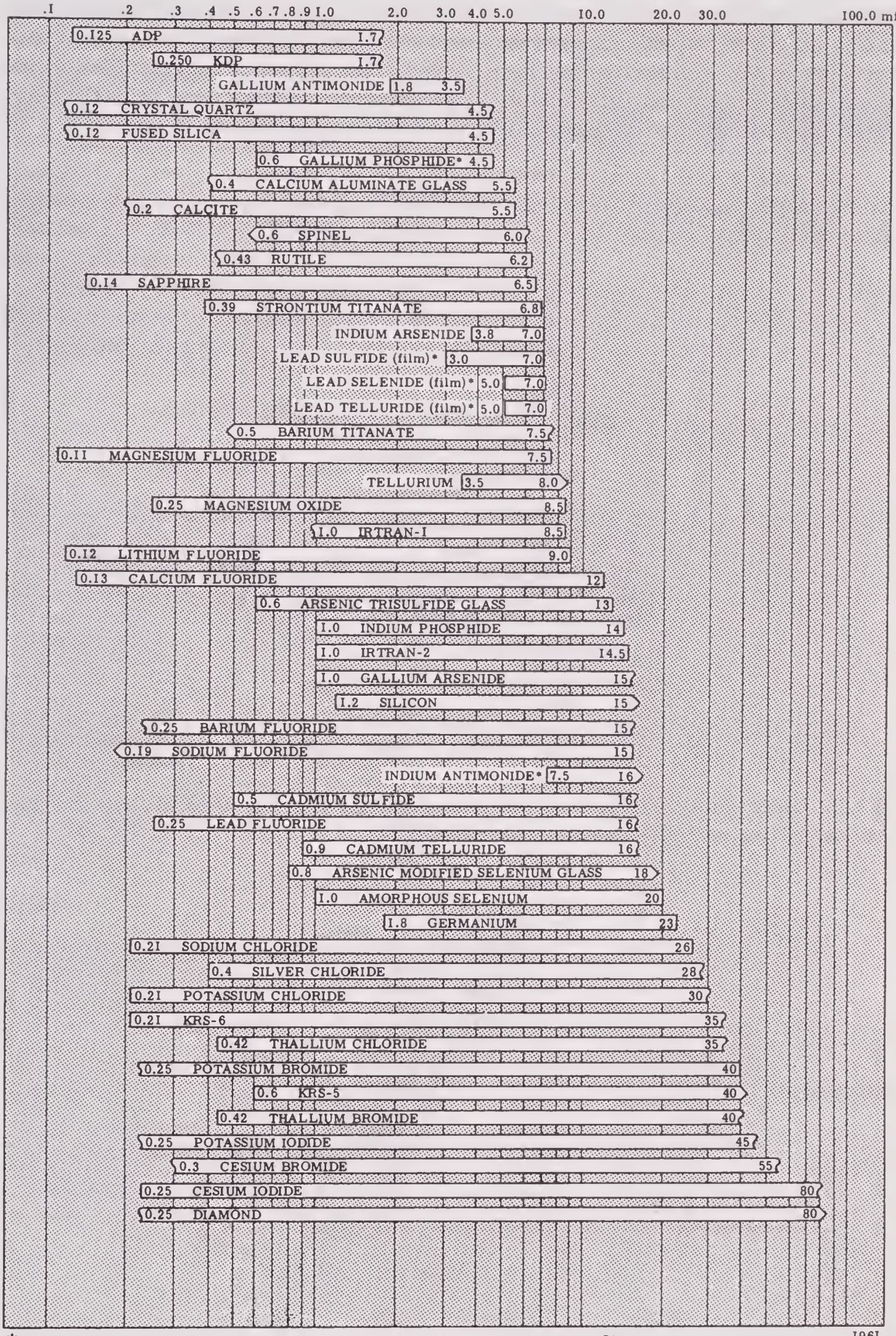
8.2.4. Some Recent Hot-Pressed Samples. Data are now becoming available on various hot-pressed materials. Irtran-1 and -2, Eastman Kodak products, are described in the supplement of [2]. Irtran-3, another Kodak product, is pressed CaF_2; it has basically the same transmission and refractive index as the single crystal. Irtran-4, pressed ZnSe, transmits to about 20 μ (0.04-in. sample transmits 50% at 21 μ) with a dip at 9 μ. Harshaw has just announced T-12, a milky white sample that is apparently hot pressed. Few data are available on its composition and properties.

8.3. Comparisons of Material Properties

8.3.1. Transmission Region. Figure 8-13 shows the transmission regions of most of the infrared optical materials. The white bars represent the wavelength region in which a particular material transmits appreciably.

The limiting wavelengths, for both high and low cutoff, are arbitrarily chosen as that wavelength at which a sample of 2-mm thickness has 10% external transmittance. For some cases this criterion is insufficient, such as in the consideration of semiconducting materials where purity and temperature must also be specified. Some semiconductors violate the criterion in another way: materials such as indium antimonide have an external transmittance of less than 10% for a 2-mm-thick sample even in their most transparent regions; these are indicated by an asterisk (*).

Several different "endings" are used for the bars of the chart in Fig. 8-13. Each has a specific meaning: a bar with a straight vertical ending indicates that the cutoff exists at the wavelength represented by the end of the bar *exactly* as defined above; a bar with an S-shape ending represents a material which cuts off at *approximately* that wavelength; a bar ending in an angle indicates that the material transmits *at least* to that wavelength, and probably further. Measurements made on materials in this last group have not been made to sufficiently long or sufficiently short wavelengths to determine the cutoff.



*Maximum external transmittance of less than 10%

FIG. 8-13. Transmission regions of optical materials, 2 mm thickness; cutoff is defined as 10 percent external transmittance, and materials marked with an asterisk never have external transmittance as high as 10 percent.

8.3.2. Refractive Index. Figure 8-14 presents curves of refractive index *versus* wavelength and illustrates the range of the refractive indices of many of the materials. For crystals whose refractive index varies with direction, only the refractive index corresponding to the ordinary ray has been plotted. The refractive indices of tellurium for both the ordinary and extraordinary rays have been omitted because of their extremely high values (approximately 6.237 and 4.789, respectively at 12 μ).

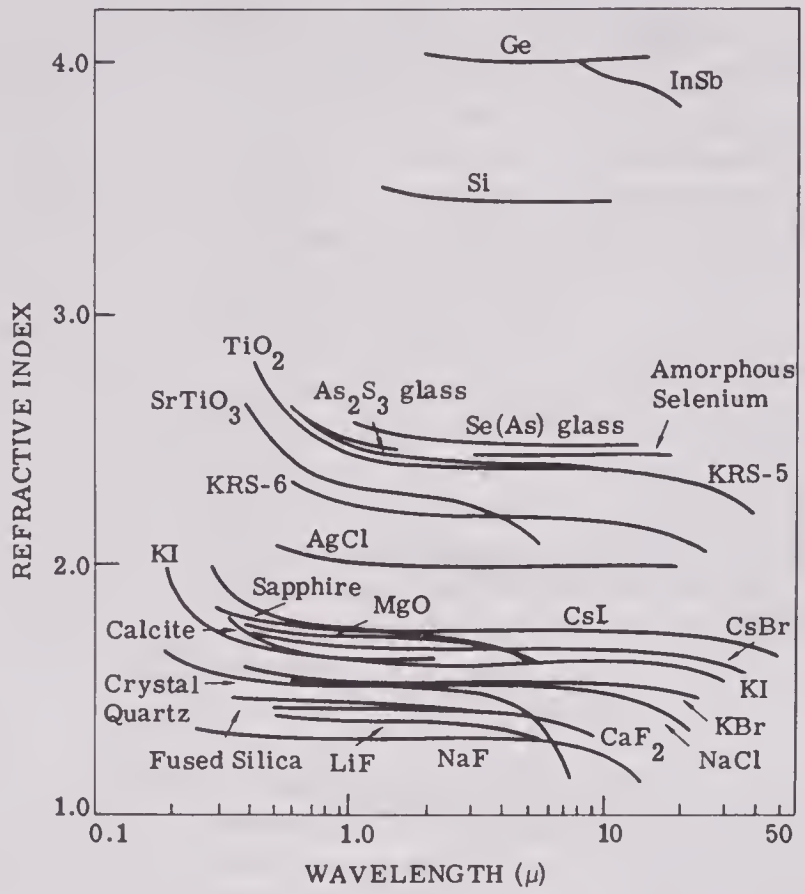


FIG. 8-14. Refractive index values.

8.3.3. Dispersion. The data of Fig. 8-14 are plotted in Fig. 8-15 to show the rate of change of the refractive index *versus* wavelength.

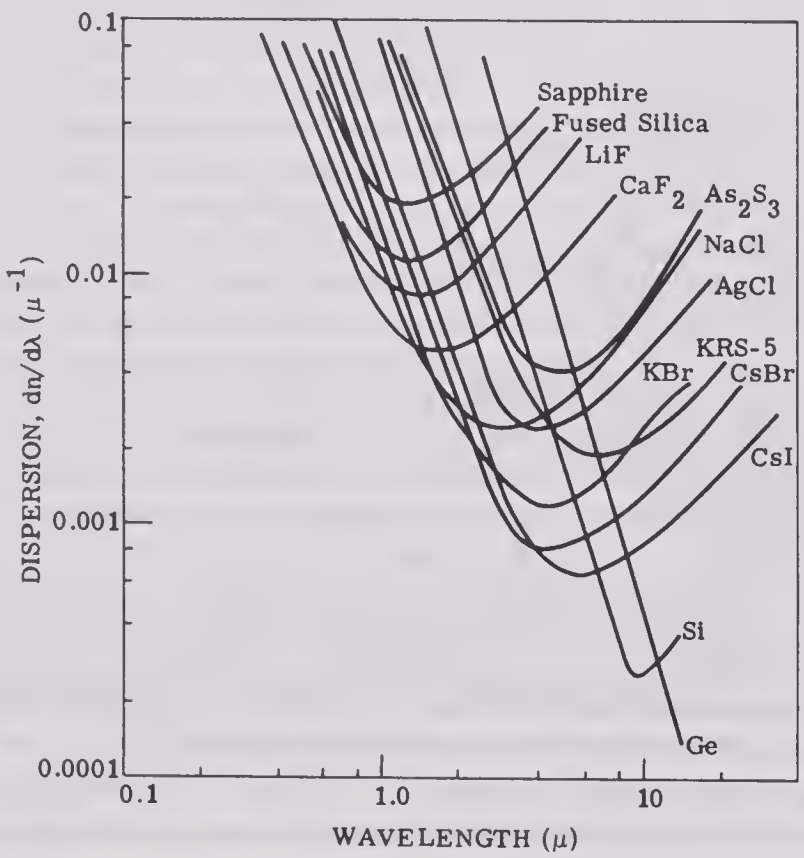


FIG. 8-15. $dn/d\lambda$ *versus* λ for selected materials.

The dispersion in the desired wavelength interval can be determined from Figs. 8-14 and 8-15. If greater accuracy is required, the dispersion can be calculated from the refractive index information contained in this chapter.

8.3.4. Dielectric Constant. Values for the dielectric constant as a function of frequency and temperature are given in Table 8-6. These values are the relative dielectric constants of materials, that is, the ratios of the dielectric constants of the material to that of a vacuum. They include measurements taken at microwave frequencies and indicate such peculiarities as variation with orientation. Since dielectric properties depend upon purity, particularly in semiconductors, the purity of the sample measured is given, where available.

When the dielectric constant is measured with the electric field parallel to the *c* axis (the optic axis), the measurement is identified with a superscript *p*; when the electric field is perpendicular to the optic axis, the measurement is identified with a superscript *s*.

TABLE 8-6. DIELECTRIC CONSTANT OF OPTICAL MATERIALS [1]

Material	*Dielectric Constant*	*Frequency* (cps)	*Temperature* (°C)	*Remarks*
Fused silica (SiO_2)	3.78	10^2 to 10^{10}	25	
Silica glass	3.81	10^8	20	
Crystal quartz (SiO_2)	4.27^p	10^7	17 to 22	
	4.34^s	10^7	17 to 22	
Potassium chloride (KCl)	4.64	10^6	29.5	
Potassium bromide (KBr)	4.90	10^2 to 10^{10}	25	
Potassium iodide (KI)	4.94	10^6	–	
Cesium iodide (CsI)	5.65	10^6	25	
Sodium chloride (NaCl)	5.90	10^2 to 10^{10}	25	
Amorphous selenium (Se)	6.00	10^2 to 10^{10}	25	
Sodium fluoride (NaF)	6.0	10^6	19	
Selenium crystal (Se)	6.0	10^2 to 10^{10}	–	
Cesium bromide (CsBr)	6.51	10^6	25	
Calcium fluoride (CaF_2)	6.76	10^5	–	
Sodium nitrate ($NaNO_3$)	6.85	10^5	19	
Mica, glass bonded, injection	6.9 to 9.2	10^6	Room	
Barium fluoride (BaF_2)	7.33	10^6	–	
Calcite ($CaCO_3$)	8.5^s	10^4	17 to 22	
	8.0^p	10^4	17 to 22	
Sapphire (Al_2O_3)	10.55^p	10^2 to 10^8	25	
	8.6^s	10^2 to 10^{10}	25	
Arsenic trisulfide glass (As_2S_3)	8.1	10^3 to 10^6	–	
Spinel ($MgO{\cdot}3.5Al_2O_3$)	8 to 9	–	–	
Lithium fluoride (LiF)	9.00	10^2 to 10^{10}	25	
Magnesium oxide (MgO)	9.65	10^2 to 10^8	25	
Cadmium telluride (CdTe)	11.0	10^3 to 10^5	–	5.5×10^{13} carriers/cc
Silicon (Si)	13	10^{10}	–	
Silver chloride (AgCl)	12.3	–	Room	
Germanium (Ge)	16.6	10^{10}	–	9.0 ohm-cm resistivity
Lead sulfide (PbS)	17.9	10^6	15	
Thallium bromide (TlBr)	30.3	10^3 to 10^7	25	
Thallium bromide-iodide (KRS-5)	32.5	10^2 to 10^7	25	
Thallium chloride (TlCl)	31.9	10^6	–	
Thallium bromide-chloride (KRS-6)	32	10^2 to 10^5	25	
Potassium dihydrogen phosphate (KDP)	44.5 to 44.3^s	10^2 to 10^8	–	
	21.4 to 20.2^p	10^2 to 10^8	–	
Ammonium dihydrogen phosphate (ADP)	56.4 to 55.9^s	10^2 to 10^8	–	
	16.4 to 13.7^p	10^2 to 10^{10}	–	
Titanium dioxide (TiO_2)	170^p	10^4 to 10^7	25	
	86^s	10^2 to 10^7	25	
Strontium titanate ($SrTiO_3$)	234	10^2 to 10^{10}	25	
Barium titanate ($BaTiO_3$)	1240 to 1200	10^2 to 10^8	25	

^{p}Dielectric constant measured parallel to *c* axis.
^{s}Dielectric constant measured perpendicular to *c* axis.
– Value not indicated.

8.3.5. Melting Temperature. The melting temperature of optical materials (or the softening temperature if appropriate, or for glasses) is given in Table 8-7.

TABLE 8-7. MELTING, OR SOFTENING, TEMPERATURE OF OPTICAL MATERIALS [1]

Material	*Temperature* (°C)
Amorphous selenium (Se)	35*
Arsenic modified selenium glass [Se(As)]	70*
Arsenic trisulfide glass (As_2S_3)	210*
Potassium dihydrogen phosphate (KDP)	252.6
Sodium nitrate ($NaNO_3$)	306.8
Gallium arsenide (GaAs)	400‡
Thallium bromide-iodide (KRS-5)	414.5
Thallium bromide-chloride (KRS-6)	423.5
Thallium chloride (TlCl)	430
Tellurium (Te)	449.7
Silver chloride (AgCl)	457.7
Thallium bromide (TlBr)	460
Gallium phosphide (GaP)	>500
Indium antimonide (InSb)	523
Cesium iodide (CsI)	621
Cesium bromide (CsBr)	636
Gallium antimonide (GaSb)	720
Potassium iodide (KI)	723
Potassium bromide (KBr)	730
Potassium chloride (KCl)	776
Sodium chloride (NaCl)	801
Borosilicate crown glass	820*
Lead fluoride (PbF_2)	855
Lithium fluoride (LiF)	870
Calcite ($CaCO_3$)	894.4‡
Cadmium sulfide (CdS)	900†
Lead telluride (PbTe)	917
Germanium (Ge)	936
Indium arsenide (InAs)	942
Sodium fluoride (NaF)	980
Cadmium telluride (CdTe)	~1040
Indium phosphide (InP)	1050
Lead selenide (PbSe)	1065
Lead sulfide (PbS)	1114
Gallium arsenide (GaAs)	1238
Barium fluoride (BaF_2)	1280
Calcium fluoride (CaF_2)	1360
Silicon (Si)	1420
Crystal quartz (SiO_2)	<1470
Barium titanate ($BaTiO_3$)	1600
Fused silica (SiO_2)	~1710
Titanium dioxide (TiO_2)	1825
Sapphire (Al_2O_3)	2030
Spinel ($MgO \cdot 3.5\, Al_2O_3$)	2030 to 2060
Strontium titanate ($SrTiO_3$)	2080
Magnesium oxide (MgO)	2800

*Softening temperature.
†Sublimation temperature.
‡Dissociation temperature.

8.3.6. Thermal Conductivity. Values of thermal conductivity for optical materials are given in Table 8-8. For crystals that exhibit anisotropy, the orientation of the heat flow with respect to the *c* axis is noted; values are given for the heat flow parallel (*p*) and perpendicular (*s*) to the *c* axis.

TABLE 8-8. THERMAL CONDUCTIVITY OF OPTICAL MATERIALS [1]

Material	*Thermal Conductivity* [10^{-4} cal/(cm sec C°)]	*Temperature* (°C)	*Remarks*
Diatomaceous earth	1.3	"ordinary"	
Arsenic modified selenium glass [Se(As)]	3.3	–	
Arsenic trisulfide glass (As_2S_3)	4.0	40	
Thallium bromide-iodide (KRS-5)	13	20	
Thallium bromide (TlBr)	14	43	
Lead sulfide (PbS)	16	–	
Thallium bromide-chloride (KRS-6)	17.1	56	
Thallium chloride (TlCl)	18	38	
Ammonium dihydrogen phosphate (ADP)	17[p]	42	
	30[s]	40	
Cesium bromide (CsBr)	23	25	
Cesium iodide (CsI)	27	25	
Silver chloride (AgCl)	27.5	22	
Fused silica (SiO_2)	28.2	41	
Potassium dihydrogen phosphate (KDP)	29[p]	39	
	32[s]	46	
Barium titanate ($BaTiO_2$)	32	Room	Ceramic material
Calcite ($CaCO_3$)	132[p]	0	
	111[s]	0	
Potassium bromide (KBr)	115	46	
Tellurium (Te)	150	–	
Sodium chloride (NaCl)	155	16	
Potassium chloride (KCl)	156	42	
Crystal quartz (SiO_2)	255[p]	50	
	148[s]	50	
Calcium fluoride (CaF_2)	232	36	
Lithium fluoride (LiF)	270	41	
Barium fluoride (BaF_2)	280	13	
Titanium dioxide (TiO_2)	300[p]	36	
	210[s]	44	
Spinel ($MgO \cdot 3.5\ Al_2O_3$)	330	35	
Cadmium sulfide (CdS)	380	20	
Sapphire (Al_2O_3)	600[p]	26	
	550[s]	23	
Magnesium oxide (MgO)	600	20	
Indium antimonide (InSb)	850	20	
Germanium (Ge)	1400	20	*n*-type, 40 ohm-cm resistivity
Silicon (Si)	3090	40	*p*-type
Silver (Ag)	10060	18	

[p]Thermal conductivity measured with heat flow parallel to *c* axis.
[s]Thermal conductivity measured with heat flow perpendicular to *c* axis.
– Value not indicated.

8.3.7. Thermal Expansion. Table 8-9 shows the linear coefficient of thermal expansion of various optical materials. For crystals exhibiting anisotropy, the orientation of heat flow with respect to the *c* axis is stated; values are given for heat flow parallel (*p*) and perpendicular (*s*) to the *c* axis.

TABLE 8-9. LINEAR COEFFICIENT OF THERMAL EXPANSION OF OPTICAL MATERIALS [1]

Material	*Coefficient of Thermal Expansion* 10^{-6}/C°	*Average Temperature or Temperature Range* (°C)	*Remarks*
Fused silica (SiO_2)	0.5	20 to 900	
Invar	0.9	20	
Silicon (Si)	4.2	25	
Cadmium sulfide (CdS)	4.2	27 to 70	
Cadmium telluride (CdTe)	4.5	50	
Indium antimonide (InSb)	4.9	20 to 60	
Indium arsenide (InAs)	5.3	–	
Germanium (Ge)	5.5 to 6.1	25	
Gallium arsenide (GaAs)	5.7	–	
Spinel ($MgO{\cdot}3.5\,Al_2O_3$)	5.9	40	
Sapphire (Al_2O_3)	6.7[p]	50	
	5.0[s]	50	
Borosilicate crown glass	9	22 to 498	
Titanium dioxide (TiO_2)	9.19[p]	40	
	7.14[s]	40	
Strontium titanate ($SrTiO_3$)	9.4	–	
Crystal quartz (SiO_2)	7.97[p]	0 to 80	
	13.37[s]	0 to 80	
Sodium nitrate ($NaNO_3$)	12[p]	50	
	11[s]	50	
Magnesium oxide (MgO)	13.8	20 to 1000	
Copper (Cu)	14.09	−191 to 16	
Tellurium (Te)	16.75	40	
Barium titanate ($BaTiO_3$)	19	10 to 70	Ceramic
	6.2[p]	4 to 20	Single crystal
	15.7[s]	4 to 20	Single crystal
Calcium fluoride (CaF_2)	24	20 to 60	
Arsenic trisulfide glass (As_2S_3)	24.6	33 to 165	
Calcite ($CaCO_3$)	25[p]	0	
	−5.8[s]	0	
Silver chloride (AgCl)	30	20 to 60	
Amorphous selenium (Se)	34	–	Estimated
Sodium fluoride (NaF)	36	Room	
Potassium chloride (KCl)	36	20 to 60	
Lithium fluoride (LiF)	37	0 to 100	
Potassium iodide (KI)	42.6	40	
Potassium bromide (KBr)	43	20 to 60	
Sodium chloride (NaCl)	44	−50 to 200	
Cesium bromide (CsBr)	47.9	20 to 50	
Thallium bromide-chloride (KRS-6)	50	20 to 100	
Cesium iodide (CsI)	50	25 to 50	
Thallium bromide (TlBr)	51	20 to 60	
Thallium chloride (TlCl)	53	20 to 60	
Thallium bromide-iodide (KRS-5)	58	20 to 100	

[p] Thermal expansion measured parallel to *c* axis.
[s] Thermal expansion measured perpendicular to *c* axis.
– Value not indicated.

8.3.8. Specific Heat. Table 8-10 lists the specific heat of optical materials. The specific heat is given at constant pressure C_p rather than at constant volume C_v, although the numerical difference is negligible for most purposes.

8.3.9. Hardness. Values of hardness for several optical materials are given in Table 8-11. Knoop values with the indenter aligned in either the ⟨100⟩ or the ⟨110⟩ direction are tabulated. The indenter load is given when it is known. Where Knoop values are not available, Moh or Vickers values are given.

8.3.10. Solubility. Values of the water solubility of optical materials at various temperatures are given in Table 8-12.

8.3.11. Young's Modulus. Values of Young's modulus for several optical materials are given in Table 8-13. The calculated values are obtained from the values of the elastic moduli by a method described in [1].

TABLE 8-10. SPECIFIC HEAT OF OPTICAL MATERIALS [1]

Material	*Specific Heat*	*Temperature* (°C)
Thallium bromide (TlBr)	0.045	20
Tellurium (Te)	0.0479	300
Cesium iodide (CsI)	0.048	20
Thallium bromide-chloride (KRS-6)	0.0482	20
Lead sulfide (PbS)	0.050	–
Thallium chloride (TlCl)	0.052	0
Cesium bromide (CsBr)	0.063	20
Germanium (Ge)	0.074	0 to 100
Potassium iodide (KI)	0.075	−3
Barium titanate ($BaTiO_3$)	0.077	−98
Silver chloride (AgCl)	0.0848	0
Potassium bromide (KBr)	0.104	0
Potassium chloride (KCl)	0.162	0
Silicon (Si)	0.168	25
Titanium dioxide (TiO_2)	0.17	25
Sapphire (Al_2O_3)	0.18	25
Crystal quartz (SiO_2)	0.188	12 to 100
Calcite ($CaCO_3$)	0.203	0
Sodium chloride (NaCl)	0.204	0
Calcium fluoride (CaF_2)	0.204	0
Magnesium oxide (MgO)	0.209	0
Fused silica (SiO_2)	0.22	–
Sodium nitrate ($NaNO_3$)	0.247	0
Sodium fluoride (NaF)	0.26	0
Lithium fluoride (LiF)	0.373	10

– Value not indicated.

Table 8-11. Hardness of Optical Materials [1]

Material	*Hardness* (Knoop Number)	*Direction*	*Indenter Load* (g)	*Remarks*
Potassium bromide (KBr)	5.9	⟨110⟩	200	
	7.0	⟨100⟩	200	
Potassium chloride (KCl)	7.2	⟨110⟩	200	
	9.3	⟨100⟩	200	
Silver chloride (AgCl)	9.5		200	
Thallium bromide (TlBr)	11.9	⟨110⟩	500	
	11.9	⟨100⟩	500	
Thallium chloride (TlCl)	12.8	⟨110⟩	500	
	12.8	⟨100⟩	500	
Sodium chloride (NaCl)	15.2	⟨110⟩	200	
	18.2	⟨100⟩	200	
Sodium nitrate ($NaNO_3$)	19.2	Perpendicular to cleavage planes	200	
Cesium bromide (CsBr)	19.5		200	
Thallium bromide-chloride (KRS-6)	29.9	⟨110⟩	500	
	38.5	⟨100⟩	500	
Thallium bromide-iodide (KRS-5)	40.2		200	Machined surface
	39.8	⟨100⟩	500	
	33.2	⟨110⟩	500	
Barium fluoride (BaF_2)	82		500	
Lithium fluoride	102-113		600	Vacuum grown
Arsenic trisulfide glass (As_2S_3)	109		100	
Calcium fluoride (CaF_2)	158.3	⟨110⟩	500	
	158.3	⟨100⟩	500	
Fused silica (SiO_2)	461		200	
Strontium titanate ($SrTiO_3$)	595			
Magnesium oxide (MgO)	692	Perpendicular to cleavage planes	600	
Crystal quartz (SiO_2)	741	Perpendicular to *z*- and *x*-cut faces	500	
Titanium dioxide (TiO_2)	879	Random	500	
Spinel ($MgO{\cdot}3.5\,Al_2O_3$)	1140	Random	1000	
Silicon (Si)	1150			7 (Moh Number)
Sapphire (Al_2O_3)	1370	Random	1000	
Calcite ($CaCO_3$)				3 (Moh Number)
Cadmium telluride (CdTe)				43.5 (Vickers Scale)
Barium titanate ($BaTiO_3$)				200-580 (Vickers Scale) Single crystal

Table 8-12. Water Solubility of Optical Materials [1]

Material	*Solubility* (g/100 g water)	*Temperature* (°C)
Spinel ($MgO{\cdot}3.5\,Al_2O_3$)	Insoluble*	
Crystal quartz (SiO_2)	Insoluble	
Fused silica (SiO_2)	Insoluble	
Titanium dioxide (TiO_2)	Insoluble	
Sapphire (Al_2O_3)	Insoluble	
Silver chloride (AgCl)	Insoluble	
Arsenic trisulfide glass (As_2S_3)	Insoluble	
Amorphous selenium (Se)	Insoluble	
Arsenic modified selenium glass [Se(As)]	Insoluble	
Tellurium (Te)	Insoluble	
Cadmium sulfide (CdS)	Insoluble	
Gallium arsenide (GaAs)	Insoluble	
Germanium (Ge)	Insoluble	
Indium arsenide (InAs)	Insoluble	
Gallium antimonide (GaSb)	Insoluble	
Tellurium (Te)	Insoluble	
Lead selenide (PbSe)	Insoluble	
Silicon (Si)	Insoluble	
Cadmium telluride (CdTe)	Insoluble	
Indium antimonide (InSb)	Insoluble	
Magnesium oxide (MgO)	Insoluble	
Calcite ($CaCO_3$)	0.0014	25
Calcium fluoride (CaF_2)	0.0017	26
Thallium bromide (TlBr)	0.05	25
Thallium bromide-iodide (KRS-5)	0.05	Room
Barium fluoride (BaF_2)	0.17	10
Lithium fluoride (LiF)	0.27	18
Thallium bromide-chloride (KRS-6)	0.32	20
Thallium chloride (TlCl)	0.32	20
Sodium fluoride (NaF)	4.22	18
Ammonium dihydrogen phosphate (ADP)	22.7	0
Potassium dihydrogen phosphate (KDP)	33	25
Potassium chloride (KCl)	34.7	20
Sodium chloride (NaCl)	35.7	0
Cesium iodide (CsI)	44	0
Potassium bromide (KBr)	53.5	0
Sodium nitrate ($NaNO_3$)	73	0
Cesium bromide (CsBr)	124.3	25
Potassium iodide (KI)	127.5	0

*"Insoluble" means less than 10^{-3} g/100 g water at room temperature.

TABLE 8-13. YOUNG'S MODULUS VALUE FOR SEVERAL OPTICAL MATERIALS [1]

Material	*Young's Modulus* (10^6 psi)	*Remarks*
Cesium iodide (CsI)	0.769	Measured in flexure
Thallium bromide-iodide (KRS-5)	2.3	Measured in flexure
Cesium bromide (CsBr)	2.3	Measured in flexure
Arsenic trisulfide glass (As_2S_3)	2.3	
Silver chloride (AgCl)	2.9	Measured in flexure
Thallium bromide-chloride (KRS-6)	3.0	Measured in flexure
Potassium bromide (KBr)	3.9	Measured in flexure
Thallium bromide (TlBr)	4.28	Calculated
Potassium chloride (KCl)	4.30	Measured in flexure
Potassium iodide (KI)	4.57	Calculated
Thallium chloride (TlCl)	4.60	Calculated
Barium titanate ($BaTiO_3$)	4.90	Single crystal
	16.50	Ceramic
Sodium chloride (NaCl)	5.80	Measured in flexure
Indium antimonide (InSb)	6.21	Calculated
Barium fluoride (BaF_2)	7.70	Measured in flexure
Gallium antimonide (GaSb)	9.19	Calculated
Lithium fluoride (LiF)	9.40	Measured in flexure Minimum value
Calcite ($CaCO_3$)	10.50[p]	
	12.80[s]	
Fused silica (SiO_2)	10.60	
Calcium fluoride (CaF_2)	11.0	Measured in flexure Minimum value
Crystal quartz (SiO_2)	11.1[s]	
	14.1[p]	
Germanium (Ge)	14.9	Calculated
Silicon (Si)	19.0	Calculated
Magnesium oxide (MgO)	36.1	Calculated
Sapphire (Al_2O_3)	50.0	

[p] Young's modulus measured parallel to *c* axis.
[s] Young's modulus measured perpendicular to *c* axis.

8.4. Material Data Useful for Lens Design

Previous sections have summarized the properties of optical materials useful for windows, domes, lenses, etc. This section includes equations that are principally useful for lens design. Every designer knows that final curvatures, etc. must be calculated for the material of an individual batch. Similarly, there are variations among crystal samples. The data given below hold for one measured sample and may be different for other samples.

8.4.1. Dispersion Equations for Individual Materials. Unfortunately, not enough data are available to compare samples, and some materials have not been of sufficient interest to warrant calculation of the constants of a dispersion equation. The available ones are listed below.

Arsenic Trisulfide Glass

$$n^2 - 1 = \sum_{i=1}^{5} \frac{K_i \lambda^2}{\lambda^2 - \lambda_i^2}$$

i	λ_i^2	K_i
1	0.0225	1.8983678
2	0.0625	1.9222979
3	0.1225	0.8765134
4	0.2025	0.1188704
5	750.	0.9569903

Cadmium Sulfide

Ordinary ray:

$$n^2 = 5.235 + \frac{1.819 \times 10^7}{\lambda^2 - 1.651 \times 10^7}$$

Extraordinary ray:

$$n^2 = 5.239 + \frac{2.076 \times 10^7}{\lambda^2 - 1.651 \times 10^7}$$

Cesium Bromide

$$n^2 = 5.640752 - 0.000003338\lambda^2 + \frac{0.0018612}{\lambda^2} + \frac{41110.49}{\lambda^2 - 14390.4} + \frac{0.0290764}{\lambda^2 - 0.024964}$$

Cesium Iodide

$$n^2 - 1 = \sum_{i=1}^{5} \frac{K_i \lambda^2}{\lambda^2 - \lambda_i^2}$$

i	λ_i	λ_i^2	K_i
1	0.34617251	0.0229567	0.00052701
2	1.0080886	0.1466	0.02.49156
3	0.28551800	0.1810	0.032761
4	0.39743178	0.2120	0.044944
5	3.3605359	161.0	25921.

Fused Silica

$$n^2 = 2.978645 + \frac{0.008777808}{\lambda^2 - 0.010609} - \frac{84.06224}{96.00000 - \lambda^2}$$

Magnesium Oxide

$$n^2 = 2.956362 - 0.1062387\lambda^2 - 0.0000204968\lambda^4 - \frac{0.02195770}{0.01428322}$$

Potassium Bromide

$$n^2 = 2.361323 - 0.00311497\lambda^2 - 0.000{,}000{,}058613\lambda^4 + \frac{0.007676}{\lambda^2} + \frac{0.0156569}{\lambda^2 - 0.0324}$$

Potassium Chloride (for the ultraviolet and visible)

$$n^2 = a^2 + \frac{M_1}{\lambda^2 - \lambda_1^2} + \frac{M_2}{\lambda^2 - \lambda_2^2} - k\lambda^2 - h\lambda^4, \quad n^2 = b^2 + \frac{M_1}{\lambda^2 - \lambda_1^2} - \frac{M_2}{\lambda^2 - \lambda_2^2} + \frac{M_3}{\lambda_3^2 - \lambda^2}$$

$a^2 = 2.174967$	$\lambda_2^2 = 0.0255550$	$b^2 = 3.866619$
$M_1 = 0.008344206$	$k = 0.000513495$	$M_3 = 5569.715$
$\lambda_1^2 = 0.0119082$	$h = 0.06167587$	$\lambda_3^2 = 3292.47$
$M_2 = 0.00698382$		

Rutile

Ordinary ray:

$$n^2 = 5.913 + 2.441 \times 10^7/(\lambda^2 - 0.803 \times 10^7)$$

Extraordinary ray:

$$n^2 = 7.197 + 3.322 \times 10^7/(\lambda^2 - 0.843 \times 10^7)$$

Silver Chloride

$$n^2 = 4.00804 - 0.00085111\lambda^2 - 0.00000019762\lambda^4 + 0.079086/(\lambda^2 - 0.04584)$$

Sphalerite

$$n^2 = 5.164 + 1.208 \times 10^7/(\lambda^2 - 0.732 \times 10^7)$$

Thallium Bromide-Iodide

$$n^2 - 1 = \sum_i \frac{K_i\lambda^2}{\lambda^2 - \lambda_i^2}$$

i	λ_i^2	K_i
1	0.0225	1.8293958
2	0.0625	1.6675593
3	0.1225	1.1210424
4	0.2025	0.04513366
5	27089.737	12.380234

8.4.2. Herzberger Dispersion Equation. Herzberger and Salzberg [4] have given an equation which uses as λ_0^2 the value 0.028 μ^2. The equation is

$$n = A + BL + CL^2 + D\lambda^2 + E\lambda^4 \tag{8-1}$$

$L = (\lambda^2 - 0.028)^{-1}$. Values for the constants A to E and the range of usefulness of the equation are given in Table 8-14. Residuals between these data and measured values are $\pm 3 \times 10^{-4}$ or less with a very few exceptions.

TABLE 8-14. CONSTANTS TO BE USED WITH THE INTERPOLATION FORMULA

No.	*Material*	*Wavelength Range* (μ) *from*	*to*	*Constant* A	B	C	D	E
1	Fused quartz	0.5	4.3	1.44902	0.004604	−0.000381	−0.0025268	−0.00007722
2	Calcium aluminate	0.6	4.3	1.64289	0.007860	−0.000231	−0.0022133	−0.00001598
3	IR-20	0.5	5.0	1.83450	0.011834	−0.000100	−0.0022268	−0.00001267
4	Strontium titanate	1.0	5.3	2.28355	0.035906	+0.001666	−0.0061335	−0.00001502
5	Magnesium oxide	0.5	5.5	1.71960	0.006305	−0.000090	−0.0031356	−0.00000770
6	Sapphire	1.0	5.6	1.75458	0.007149	−0.001577	−0.0045380	−0.00002808
7	Lithium fluoride	0.5	6.0	1.38761	0.001796	−0.000041	−0.0023045	−0.00000557
8	Irtran 1	1.0	6.7	1.37770	0.001348	+0.000216	−0.0015041	−0.00000441
9	Calcium fluoride	0.6	8.3	1.42780	0.002267	−0.000069	−0.0011157	−0.00000162
10	Barium fluoride	0.5	11.0	1.46629	0.002867	+0.000064	−0.0006035	−0.000000465
11	Silicon	1.3	11.0	3.41696	0.138497	+0.013924	−0.0000209	+0.000000148
12	Arsenic trisulfide	0.6	12.0	2.41326	0.055720	+0.006177	−0.0003044	−0.000000232
13	Irtran 2	1.0	13.5	2.25698	0.032586	+0.000679	−0.0005272	−0.000000604
14	Germanium	2.0	13.5	3.99931	0.391707	+0.163492	−0.0000060	+0.000000053

8.4.3. Refractive Index, Dispersion Values. Figure 8-2 shows the usual "glass table." From such a plot the optical designer can choose glasses to balance powers and chromatic aberrations. For infrared lens design, a different Abbé number (abscissa of Fig. 8-2) must be defined. Three definitions appear to be useful:

$$\nu_{1\text{-}2.7} = \frac{n_{1.85} - 1}{n_{2.7} - n_{1.0}}$$

$$\nu_{3\text{-}5.5} = \frac{n_{4.25} - 1}{n_{5.5} - 1} \qquad (8\text{-}2)$$

$$\nu_{8\text{-}15} = \frac{n_{11.5} - 1}{n_{15} - n_8}$$

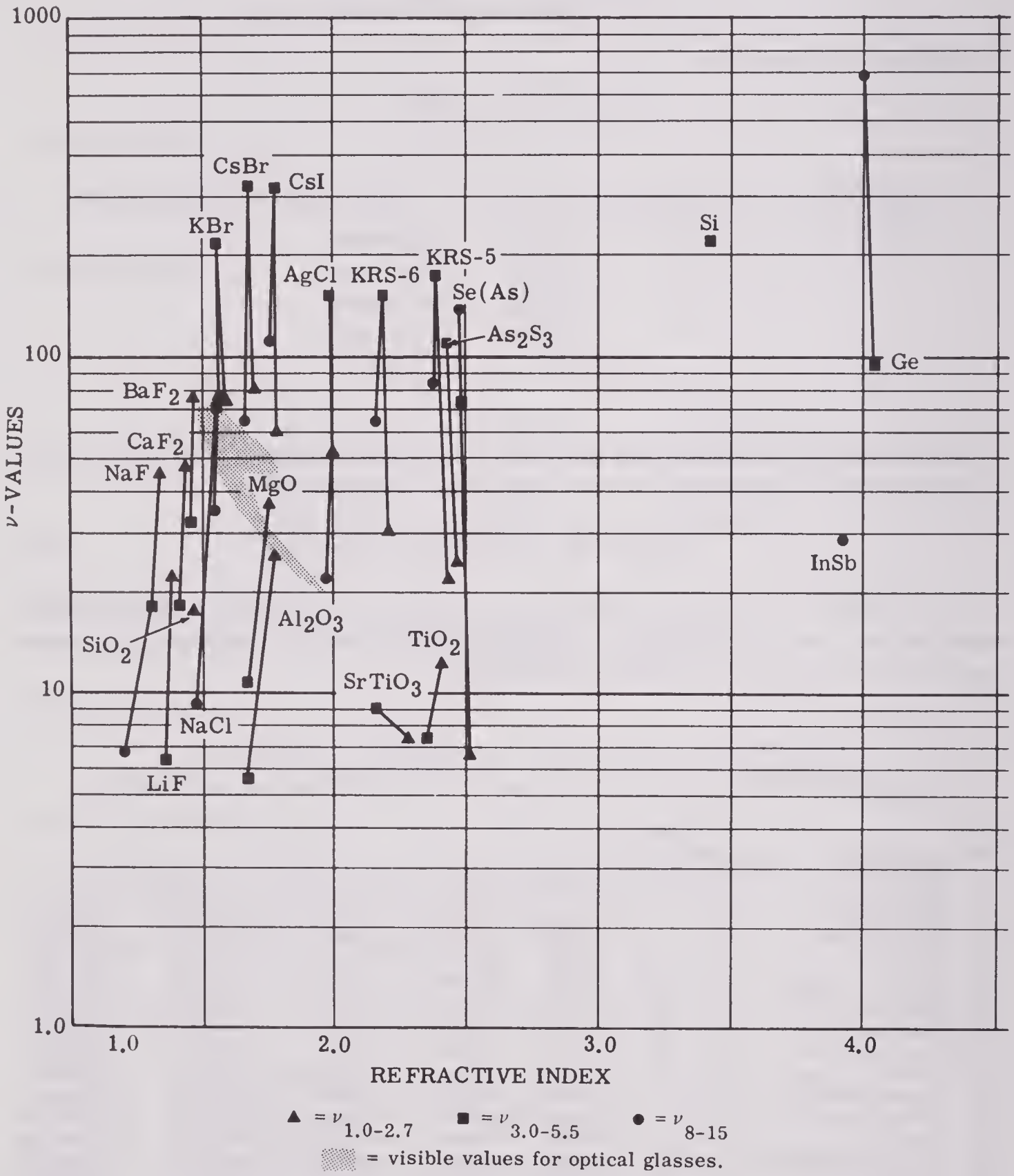


FIG. 8-16. Reciprocal dispersion of some infrared optical materials.

Values are plotted in Fig. 8-16. The lines connect values of ν in different spectral regions for the same material.

Values of partial dispersion are also useful. They can be easily calculated from index data. A curve for 15 materials useful from 8 μ to 15 μ is given in Fig. 8-17. Such a curve, which is easily constructed for other materials, is useful in the design of highly color-corrected triplet lenses (P is the partial dispersion). Equations for the combined power of a system of thin lenses (Chapter 9) show that the differences in ν of the first and last elements should be large, and that the change in $P\,\Delta P$ should be large. In this case ΔP is the difference between the line connecting the curves for two materials and the third material.

Another useful technique invented by Szeles and Cuny [5] employs the concept of angular dispersion. Quantities are defined in Fig. 8-18, which also gives data for a number of useful materials.

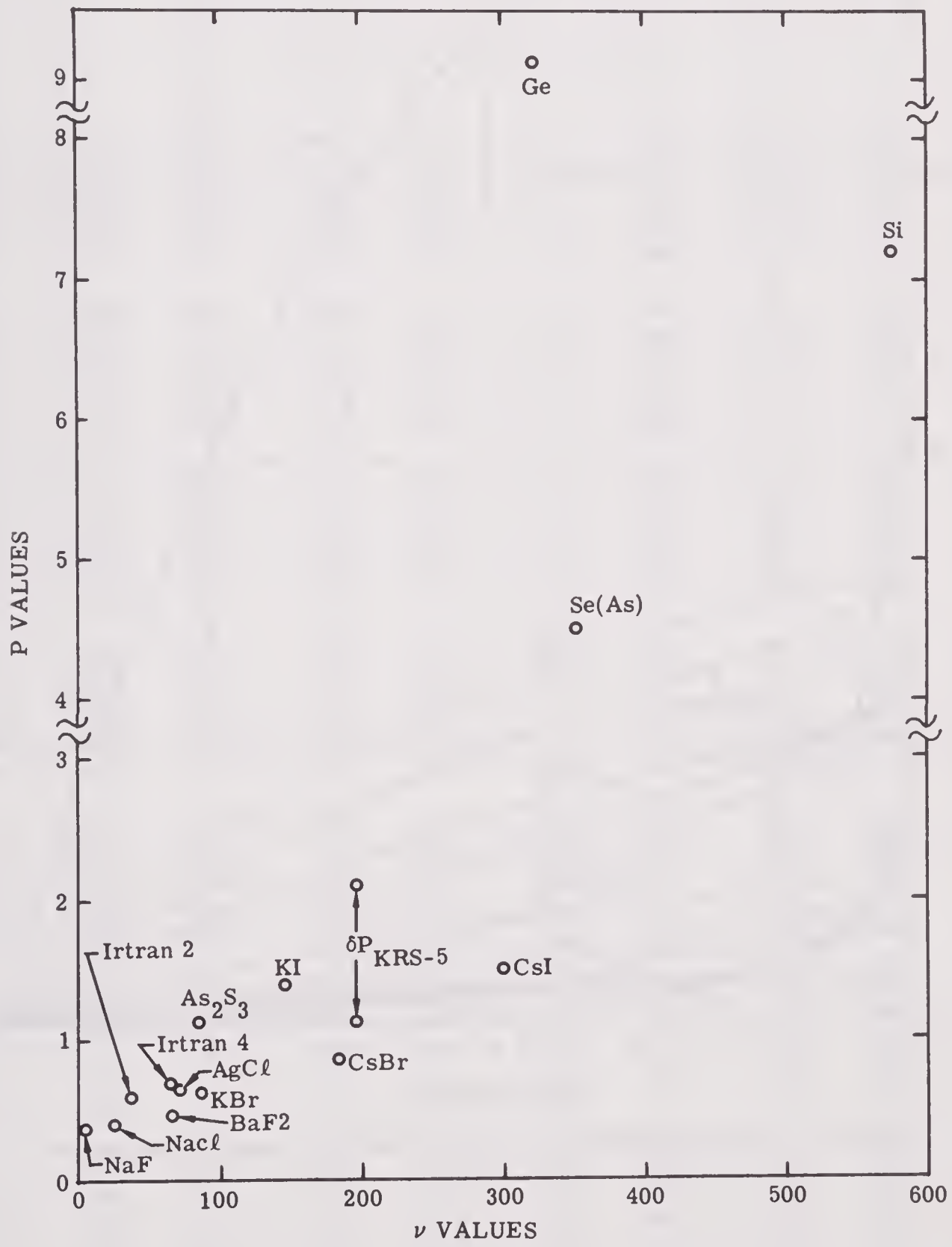


FIG. 8-17. Dispersive power of 15 materials for use in the infrared.

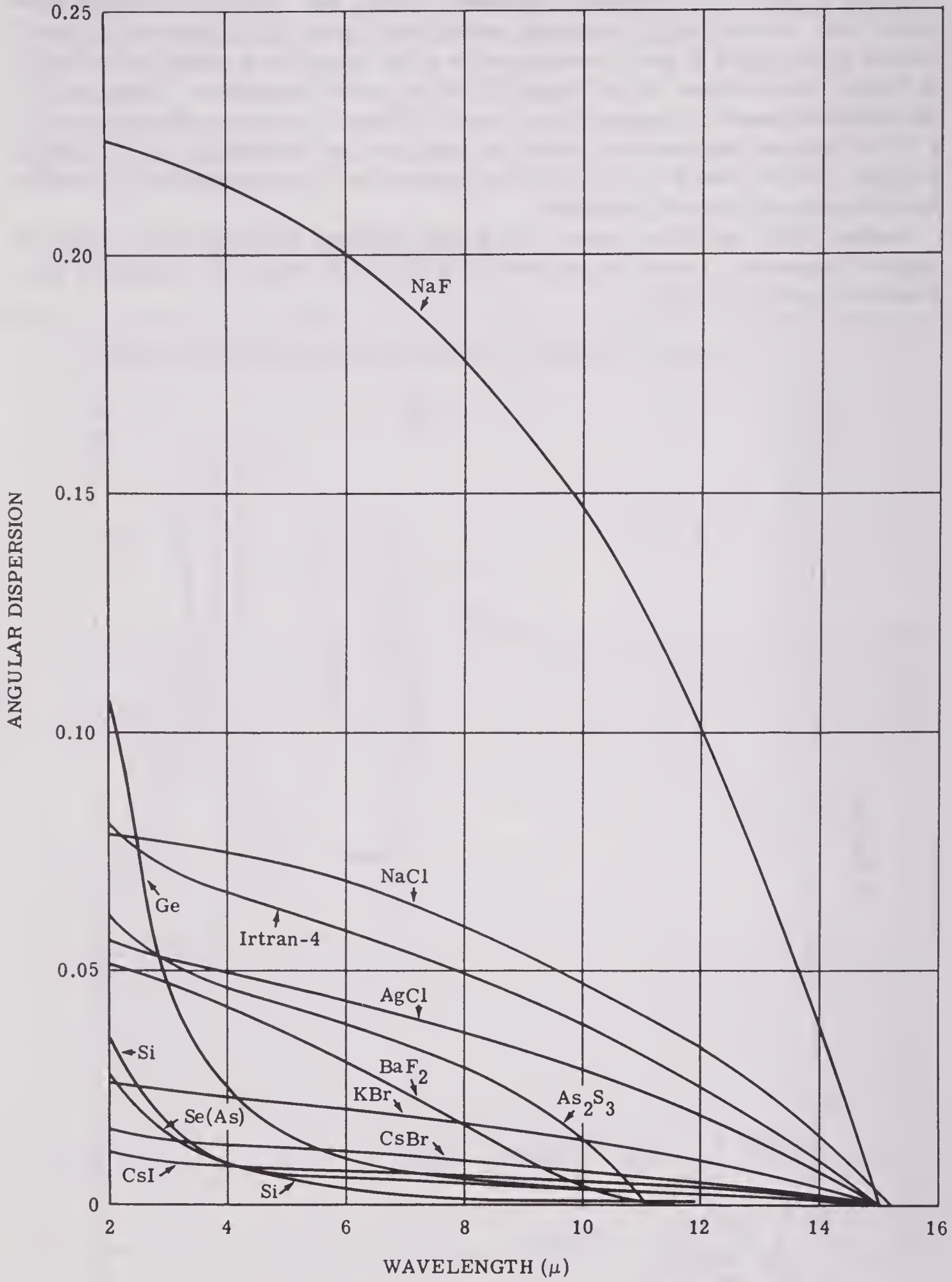


FIG. 8-18. The angular dispersion of a thin lens as a function of wavelength.

8.4.4. G Sums. In lens design, Conrady [6] defines and uses terms called G sums. Values for the sums are listed in Table 8-15.

$$G_0 = \frac{1}{2}(n-1) \qquad G_5 = 4\left(\frac{n+1}{n}\right) G_0$$

$$G_1 = n^2 G_0 \qquad G_6 = \left(\frac{3n+2}{n}\right) G_0$$

$$G_2 = (2n+1) G_0 \qquad G_7 = \left(\frac{2n+1}{n}\right) G_0$$

$$G_3 = (3n+1) G_0 \qquad G_8 = nG_0$$

$$G_4 = \left(\frac{n+2}{n}\right) G_0 \qquad G_9 = \frac{G_0}{n}$$

TABLE 8-15. CONRADY G SUMS

N	G_1	G_2	G_3	G_4	G_5	G_6	G_7	G_8	G_9
1.3500	0.3189	0.6475	0.8837	0.4343	1.2185	0.7843	0.4796	0.2362	0.1296
.3600	.3329	.6696	.9144	.4447	.2494	.8047	.4924	.2448	.1324
.3700	.3472	.6919	.9453	.4551	.2801	.8251	.5050	.2534	.1350
.3800	.3618	.7144	.9766	.4654	.3107	.8454	.5177	.2622	.1377
.3900	.3768	.7371	1.0081	.4756	.3412	.8656	.5303	.2710	.1403
.4000	.3920	.7600	.0400	.4857	.3714	.8857	.5429	.2800	.1429
.4100	.4076	.7831	.0721	.4958	.4016	.9058	.5554	.2890	.1454
.4200	.4234	.8064	.1046	.5058	.4315	.9258	.5679	.2982	.1479
.4300	.4397	.8299	.1373	.5157	.4614	.9457	.5803	.3074	.1503
.4400	.4562	.8536	.1704	.5256	.4911	.9656	.5928	.3168	.1528
.4500	.4731	.8775	.2037	.5353	.5207	.9853	.6052	.3262	.1552
.4600	.4903	.9016	.2374	.5451	.5501	1.0051	.6175	.3358	.1575
.4700	.5078	.9259	.2713	.5547	.5795	.0247	.6299	.3454	.1599
.4800	.5257	.9504	.3056	.5643	.6086	.0443	.6422	.3552	.1622
.4900	.5439	.9751	.3401	.5739	.6377	.0639	.6544	.3650	.1644
.5000	.5625	1.0000	.3750	.5833	.6667	.0833	.6667	.3750	.1667
.5100	.5814	.0251	.4101	.5927	.6955	.1027	.6789	.3850	.1689
.5200	.6007	.0504	.4456	.6021	.7242	.1221	.6911	.3952	.1711
.5300	.6203	.0759	.4813	.6114	.7528	.1414	.7032	.4054	.1732
.5400	.6403	.1016	.5174	.6207	.7813	.1606	.7153	.4158	.1753
.5500	.6607	.1275	.5537	.6298	.8097	.1798	.7274	.4262	.1774
.5600	.6814	.1536	.5904	.6390	.8379	.1990	.7395	.4368	.1795
.5700	.7025	.1799	.6273	.6481	.8661	.2181	.7515	.4474	.1815
.5800	.7240	.2064	.6646	.6571	.8942	.2371	.7635	.4582	.1835
.5900	.7458	.2331	.7021	.6661	.9221	.2561	.7755	.4690	.1855
.6000	.7680	.2600	.7400	.6750	.9500	.2750	.7875	.4800	.1875
.6100	.7906	.2871	.7781	.6839	.9778	.2939	.7994	.4910	.1894
.6200	.8136	.3144	.8166	.6927	2.0054	.3127	.8114	.5022	.1914
.6300	.8369	.3419	.8553	.7015	.0330	.3315	.8233	.5134	.1933
.6400	.8607	.3696	.8944	.7102	.0605	.3502	.8351	.5248	.1951
.6500	.8848	.3975	.9337	.7189	.0879	.3689	.8470	.5362	.1970
.6600	.9093	.4256	.9734	.7276	.1152	.3876	.8588	.5478	.1988
.6700	.9343	.4539	2.0133	.7362	.1424	.4062	.8706	.5594	.2006
.6800	.9596	.4824	.0536	.7448	.1695	.4248	.8824	.5712	.2024
.6900	.9854	.5111	.0941	.7533	.1966	.4433	.8941	.5830	.2041

TABLE 8-15. CONRADY G SUMS (*Continued*)

N	G_1	G_2	G_3	G_4	G_5	G_6	G_7	G_8	G_9
1.7000	1.0115	1.5400	2.1350	0.7618	2.2235	1.4618	0.9059	0.5950	0.2059
.7100	.0381	.5691	.1761	.7702	.2504	.4802	.9176	.6070	.2076
.7200	.0650	.5984	.2176	.7786	.2772	.4986	.9293	.6192	.2093
.7300	.0924	.6279	.2593	.7870	.3039	.5170	.9410	.6314	.2110
.7400	.1202	.6576	.3014	.7953	.3306	.5353	.9526	.6438	.2126
.7500	.1484	.6875	.3437	.8036	.3571	.5536	.9643	.6562	.2143
.7600	.1771	.7176	.3864	.8118	.3836	.5718	.9759	.6688	.2159
.7700	.2062	.7479	.4293	.8200	.4101	.5900	.9875	.6814	.2175
.7800	.2357	.7784	.4726	.8282	.4364	.6082	.9991	.6942	.2191
.7900	.2656	.8091	.5161	.8363	.4627	.6263	1.0107	.7070	.2207
.8000	.2960	.8400	.5600	.8444	.4889	.6444	.0222	.7200	.2222
.8100	.3268	.8711	.6041	.8525	.5150	.6625	.0338	.7330	.2238
.8200	.3581	.9024	.6486	.8605	.5411	.6805	.0453	.7462	.2253
.8300	.3898	.9339	.6933	.8686	.5671	.6986	.0568	.7594	.2268
.8400	.4220	.9656	.7384	.8765	.5930	.7165	.0683	.7728	.2283
.8500	.4546	.9975	.7837	.8845	.6189	.7345	.0797	.7862	.2297
.8600	.4876	2.0296	.8294	.8924	.6447	.7524	.0912	.7998	.2312
.8700	.5212	.0619	.8753	.9002	.6705	.7702	.1026	.8134	.2326
.8800	.5551	.0944	.9216	.9081	.6962	.7881	.1140	.8272	.2340
.8900	.5896	.1271	.9681	.9159	.7218	.8059	.1254	.8410	.2354
.9000	.6245	.1600	3.0150	.9237	.7474	.8237	.1368	.8550	.2368
.9100	.6599	.1931	.0621	.9314	.7729	.8414	.1482	.8690	.2382
.9200	.6957	.2264	.1096	.9392	.7983	.8592	.1596	.8832	.2396
.9300	.7321	.2599	.1573	.9469	.8237	.8769	.1709	.8974	.2409
.9400	.7689	.2936	.2054	.9545	.8491	.8945	.1823	.9118	.2423
.9500	.8062	.3275	.2537	.9622	.8744	.9122	.1936	.9262	.2436
.9600	.8440	.3616	.3024	.9698	.8996	.9298	.2049	.9408	.2449
.9700	.8822	.3959	.3513	.9774	.9248	.9474	.2162	.9554	.2462
.9800	.9210	.4304	.4006	.9849	.9499	.9649	.2275	.9702	.2475
.9900	.9602	.4651	.4501	.9925	.9750	.9825	.2387	.9850	.2487
2.0000	2.0000	.5000	.5000	1.0000	3.0000	2.0000	.2500	1.0000	.2500
.0100	.0403	.5351	.5501	.0075	.0250	.0175	.2612	.0150	.2512
.0200	.0810	.5704	.6006	.0150	.0499	.0350	.2725	.0302	.2525
.0300	.1223	.6059	.6513	.0224	.0748	.0524	.2837	.0454	.2537
.0400	.1640	.6416	.7024	.0298	.0996	.0698	.2949	.0608	.2549
.0500	.2063	.6775	.7537	.0372	.1244	.0872	.3061	.0762	.2561
.0600	.2491	.7136	.8054	.0446	.1491	.1046	.3173	.0918	.2573
.0700	.2924	.7499	.8573	.0519	.1738	.1219	.3285	.1074	.2585
.0800	.3363	.7864	.9096	.0592	.1985	.1392	.3396	.1232	.2596
.0900	.3806	.8231	.9621	.0665	.2231	.1565	.3508	.1390	.2608
.1000	.4255	.8600	4.0150	.0738	.2476	.1738	.3619	.1550	.2619
.1100	.4709	.8971	.0681	.0811	.2721	.1911	.3730	.1710	.2630
.1200	.5169	.9344	.1216	.0883	.2966	.2083	.3842	.1872	.2642
.1300	.5633	.9719	.1753	.0955	.3210	.2255	.3953	.2034	.2653
.1400	.6104	3.0096	.2294	.1027	.3454	.2427	.4064	.2198	.2664
.1500	.6579	.0475	.2837	.1099	.3698	.2599	.4174	.2362	.2674
.1600	.7060	.0856	.3384	.1170	.3941	.2770	.4285	.2528	.2685
.1700	.7547	.1239	.3933	.1242	.4183	.2942	.4396	.2694	.2696
.1800	.8039	.1624	.4486	.1313	.4426	.3113	.4506	.2862	.2706
.1900	.8537	.2011	.5041	.1384	.4668	.3284	.4617	.3030	.2717

TABLE 8-15. CONRADY G SUMS (*Continued*)

N	G_1	G_2	G_3	G_4	G_5	G_6	G_7	G_8	G_9
2.2000	2.9040	3.2400	4.5600	1.1455	3.4909	2.3455	1.4727	1.3200	0.2727
.2100	.9549	.2791	.6161	.1525	.5150	.3625	.4838	.3370	.2738
.2200	3.0063	.3184	.6726	.1595	.5391	.3795	.4948	.3542	.2748
.2300	.0583	.3579	.7293	.1666	.5631	.3966	.5058	.3714	.2758
.2400	.1109	.3976	.7864	.1736	.5871	.4136	.5168	.3888	.2768
.2500	.1641	.4375	.8437	.1806	.6111	.4306	.5278	.4062	.2778
.2600	.2178	.4776	.9014	.1875	.6350	.4475	.5388	.4238	.2788
.2700	.2721	.5179	.9593	.1945	.6589	.4645	.5497	.4414	.2797
.2800	.3270	.5584	5.0176	.2014	.6828	.4814	.5607	.4592	.2807
.2900	.3824	.5991	.0761	.2083	.7066	.4983	.5717	.4770	.2817
.3000	.4385	.6400	.1350	.2152	.7304	.5152	.5826	.4950	.2826
.3100	.4951	.6811	.1941	.2221	.7542	.5321	.5935	.5130	.2835
.3200	.5524	.7224	.2536	.2290	.7779	.5490	.6045	.5312	.2845
.3300	.6102	.7639	.3133	.2358	.8016	.5658	.6145	.5494	.2854
.3400	.6687	.8056	.3734	.2426	.8253	.5826	.6263	.5678	.2863
.3500	.7277	.8475	.4337	.2495	.8489	.5995	.6372	.5862	.2872
.3600	.7873	.8896	.4944	.2563	.8725	.6163	.6481	.6048	.2881
.3700	.8476	.9319	.5553	.2631	.8961	.6331	.6590	.6234	.2890
.3800	.9084	.9744	.6166	.2698	.9197	.6498	.6699	.6422	.2899
.3900	.9699	4.0171	.6781	.2766	.9432	.6666	.6808	.6610	.2908
.4000	4.0320	.0600	.7400	.2833	.9667	.6833	.6917	.6800	.2917
.4100	.0947	.1031	.8021	.2901	.9901	.7001	.7025	.6990	.2925
.4200	.1580	.1464	.8646	.2968	4.0136	.7168	.7134	.7182	.2934
.4300	.2220	.1899	.9273	.3035	.0370	.7335	.7242	.7374	.2942
.4400	.2866	.2336	.9904	.3102	.0603	.7502	.7351	.7568	.2951
.4500	.3518	.2775	6.0537	.3168	.0837	.7668	.7459	.7762	.2959
.4600	.4177	.3216	.1174	.3235	.1070	.7835	.7567	.7958	.2967
.4700	.4842	.3659	.1813	.3301	.1303	.8001	.7676	.8154	.2976
.4800	.5513	.4104	.2456	.3368	.1535	.8168	.7784	.8352	.2984
.4900	.6191	.4551	.3101	.3434	.1768	.8334	.7892	.8550	.2992
.5000	.6875	.5000	.3750	.3500	.2000	.8500	.8000	.8750	.3000
.5100	.7566	.5451	.4401	.3566	.2232	.8666	.8108	.8950	.3008
.5200	.8263	.5904	.5056	.3632	.2463	.8832	.8216	.9152	.3016
.5300	.8967	.6359	.5713	.3697	.2695	.8997	.8324	.9354	.3024
.5400	.9677	.6816	.6374	.3763	.2926	.9163	.8431	.9558	.3031
.5500	5.0394	.7275	.7037	.3828	.3157	.9328	.8539	.9762	.3039
.5600	.1118	.7736	.7704	.3894	.3387	.9494	.8647	.9968	.3047
.5700	.1848	.8199	.8373	.3959	.3618	.9659	.8754	2.0174	.3054
.5800	.2586	.8664	.9046	.4024	.3848	.9824	.8862	.0382	.3062
.5900	.3329	.9131	.9721	.4089	.4078	.9989	.8969	.0590	.3069
.6000	.4080	.9600	7.0400	.4154	.4308	3.0154	.9077	.0800	.3077
.6100	.4837	5.0071	.1081	.4219	.4537	.0319	.9184	.1010	.3084
.6200	.5602	.0544	.1766	.4283	.4766	.0483	.9292	.1222	.3092
.6300	.6373	.1019	.2453	.4348	.4995	.0648	.9399	.1434	.3099
.6400	.7151	.1496	.3144	.4412	.5224	.0812	.9506	.1648	.3106
.6500	.7936	.1975	.3837	.4476	.5453	.0976	.9613	.1862	.3113
.6600	.8727	.2456	.4534	.4541	.5681	.1141	.9720	.2078	.3120
.6700	.9526	.2939	.5233	.4605	.5909	.1305	.9827	.2294	.3127
.6800	6.0332	.3424	.5936	.4669	.6137	.1469	.9934	.2512	.3134
.6900	.1145	.3911	.6641	.4733	.6365	.1633	2.0041	.2730	.3141

TABLE 8-15. CONRADY G SUMS (*Continued*)

N	G_1	G_2	G_3	G_4	G_5	G_6	G_7	G_8	G_9
2.7000	6.1965	5.4400	7.7350	1.4796	4.6593	3.1796	2.0148	2.2950	0.3148
.7100	.2792	.4891	.8061	.4860	.6820	.1960	.0255	.3170	.3155
.7200	.3626	.5384	.8776	.4924	.7047	.2124	.0362	.3392	.3162
.7300	.4468	.5879	.9493	.4987	.7274	.2287	.0468	.3614	.3168
.7400	.5316	.6376	8.0214	.5050	.7501	.2450	.0575	.3838	.3175
.7500	.6172	.6875	.0937	.5114	.7727	.2614	.0682	.4062	.3182
.7600	.7035	.7376	.1664	.5177	.7954	.2777	.0788	.4288	.3188
.7700	.7905	.7879	.2393	.5240	.8180	.2940	.0895	.4514	.3195
.7800	.8783	.8384	.3126	.5303	.8406	.3103	.1001	.4742	.3201
.7900	.9668	.8891	.3861	.5366	.8632	.3266	.1108	.4970	.3208
.8000	7.0560	.9400	.4600	.5429	.8857	.3429	.1214	.5200	.3214
.8100	.1460	.9911	.5341	.5491	.9083	.3591	.1321	.5430	.3221
.8200	.2367	6.0424	.6086	.5554	.9308	.3754	.1427	.5662	.3227
.8300	.3281	.0939	.6833	.5616	.9533	.3916	.1533	.5894	.3233
.8400	.4204	.1456	.7584	.5679	.9758	.4079	.1639	.6128	.3239
.8500	.5133	.1975	.8337	.5741	.9982	.4241	.1746	.6362	.3246
.8600	.6070	.2496	.9094	.5803	5.0207	.4403	.1852	.6598	.3252
.8700	.7015	.3019	.9853	.5866	.0431	.4566	.1958	.6834	.3258
.8800	.7967	.3544	9.0616	.5928	.0656	.4728	.2064	.7072	.3264
.8900	.8927	.4071	.1381	.5990	.0880	.4890	.2170	.7310	.3270
.9000	.9895	.4600	.2150	.6052	.1103	.5052	.2276	.7550	.3276
.9100	8.0870	.5131	.2921	.6114	.1327	.5214	.2382	.7790	.3282
.9200	.1853	.5664	.3696	.6175	.1551	.5375	.2488	.8032	.3288
.9300	.2844	.6199	.4473	.6237	.1774	.5537	.2594	.8274	.3294
.9400	.3843	.6736	.5254	.6299	.1997	.5699	.2699	.8518	.3299
.9500	.4849	.7275	.6037	.6360	.2220	.5860	.2805	.8762	.3305
.9600	.5864	.7816	.6824	.6422	.2443	.6022	.2911	.9008	.3311
.9700	.6886	.8359	.7613	.6483	.2666	.6183	.3016	.9254	.3316
.9800	.7916	.8904	.8406	.6544	.2889	.6344	.3122	.9502	.3322
.9900	.8954	.9451	.9201	.6606	.3111	.6506	.3228	.9750	.3328
3.0000	9.0000	7.0000	10.0000	.6667	.3333	.6667	.3333	3.0000	.3333
.0100	.1054	.0551	.0801	.6728	.3555	.6828	.3439	.0250	.3339
.0200	.2116	.1104	.1606	.6789	.3777	.6989	.3544	.0502	.3344
.0300	.3186	.1659	.2413	.6850	.3999	.7150	.3650	.0754	.3350
.0400	.4264	.2216	.3224	.6911	.4221	.7311	.3755	.1008	.3355
.0500	.5351	.2775	.4037	.6971	.4443	.7471	.3861	.1262	.3361
.0600	.6445	.3336	.4854	.7032	.4664	.7632	.3966	.1518	.3366
.0700	.7548	.3899	.5673	.7093	.4885	.7793	.4071	.1774	.3371
.0800	.8659	.4464	.6496	.7153	.5106	.7953	.4177	.2032	.3377
.0900	.9778	.5031	.7321	.7214	.5328	.8114	.4282	.2290	.3382
.1000	10.0905	.5600	.8150	.7274	.5548	.8274	.4387	.2550	.3387
.1100	.2041	.6171	.8981	.7335	.5769	.8435	.4492	.2810	.3392
.1200	.3185	.6744	.9816	.7395	.5990	.8595	.4597	.3072	.3397
.1300	.4337	.7319	11.0653	.7455	.6210	.8755	.4703	.3334	.3403
.1400	.5498	.7896	.1494	.7515	.6431	.8915	.4808	.3598	.3408
.1500	.6667	.8475	.2337	.7575	.6651	.9075	.4913	.3862	.3413
.1600	.7844	.9056	.3184	.7635	.6871	.9235	.5018	.4128	.3418
.1700	.9031	.9639	.4033	.7695	.7091	.9395	.5123	.4394	.3423
.1800	11.0225	8.0224	.4886	.7755	.7311	.9555	.5228	.4662	.3428
.1900	.1428	.0811	.5741	.7815	.7530	.9715	.5333	.4930	.3433

TABLE 8-15. CONRADY G SUMS (*Continued*)

N	G_1	G_2	G_3	G_4	G_5	G_6	G_7	G_8	G_9
3.2000	11.2640	8.1400	11.6600	1.7875	5.7750	3.9875	2.5437	3.5200	0.3437
.2100	.3860	.1991	.7461	.7935	.7969	4.0035	.5542	.5470	.3442
.2200	.5089	.2584	.8326	.7994	.8189	.0194	.5647	.5742	.3447
.2300	.6327	.3179	.9193	.8054	.8408	.0354	.5752	.6014	.3452
.2400	.7573	.3776	12.0064	.8114	.8627	.0514	.5857	.6288	.3457
.2500	.8828	.4375	.0937	.8173	.8846	.0673	.5962	.6562	.3462
.2600	12.0092	.4976	.1814	.8233	.9065	.0833	.6066	.6838	.3466
.2700	.1364	.5579	.2693	.8292	.9284	.0992	.6171	.7114	.3471
.2800	.2646	.6184	.3576	.8351	.9502	.1151	.6276	.7392	.3476
.2900	.3936	.6791	.4461	.8410	.9721	.1310	.6380	.7670	.3480
.3000	.5235	.7400	.5350	.8470	.9939	.1470	.6485	.7950	.3485
.3100	.6543	.8011	.6241	.8529	6.0158	.1629	.6589	.8230	.3489
.3200	.7860	.8624	.7136	.8588	.0376	.1788	.6694	.8512	.3494
.3300	.9186	.9239	.8033	.8647	.0594	.1947	.6798	.8794	.3498
.3400	.0521	.9856	.8934	.8706	.0812	.2106	.6903	.9078	.3503
.3500	13.1864	9.0475	.9837	.8765	.1030	.2265	.7007	.9362	.3507
.3600	.3217	.1096	13.0744	.8824	.1248	.2424	.7112	.9648	.3512
.3700	.4579	.1719	.1653	.8883	.1465	.2583	.7216	.9934	.3516
.3800	.5950	.2344	.2566	.8941	.1683	.2741	.7321	4.0222	.3521
.3900	.7331	.2971	.3481	.9000	.1900	.2900	.7425	.0510	.3525
.4000	.8720	.3600	.4400	.9059	.2118	.3059	.7529	.0800	.3529
.4100	14.0119	.4231	.5321	.9117	.2335	.3217	.7634	.1090	.3534
.4200	.1526	.4864	.6246	.9176	.2552	.3376	.7738	.1382	.3538
.4300	.2944	.5499	.7173	.9235	.2769	.3535	.7842	.1674	.3542
.4400	.4370	.6136	.8104	.9293	.2986	.3693	.7947	.1968	.3547
.4500	.5806	.6775	.9037	.9351	.3203	.3851	.8051	.2262	.3551
.4600	.7251	.7416	.9974	.9410	.3420	.4010	.8155	.2558	.3555
.4700	.8705	.8059	14.0913	.9468	.3636	.4168	.8259	.2854	.3559
.4800	15.0169	.8704	.1856	.9526	.3853	.4326	.8363	.3152	.3563
.4900	.1642	.9351	.2801	.9585	.4069	.4485	.8467	.3450	.3567
.5000	.3125	10.0000	.3750	.9643	.4286	.4643	.8571	.3750	.3571
.5100	.4617	.0651	.4701	.9701	.4502	.4801	.8675	.4050	.3575
.5200	.6119	.1304	.5656	.9759	.4718	.4959	.8780	.4352	.3580
.5300	.7630	.1959	.6613	.9817	.4934	.5117	.8884	.4654	.3584
.5400	.9151	.2616	.7574	.9875	.5150	.5275	.8988	.4958	.3588
.5500	16.0682	.3275	.8537	.9933	.5366	.5433	.9092	.5262	.3592
.5600	.2222	.3936	.9504	.9991	.5582	.5591	.9196	.5568	.3596
.5700	.3772	.4599	15.0473	2.0049	.5798	.5749	.9299	.5874	.3599
.5800	.5332	.5264	.1446	.0107	.6013	.5907	.9403	.6182	.3603
.5900	.6901	.5931	.2421	.0164	.6229	.6064	.9507	.6490	.3607
.6000	.8480	.6600	.3400	.0222	.6444	.6222	.9611	.6800	.3611
.6100	17.0069	.7271	.4381	.0280	.6660	.6380	.9715	.7110	.3615
.6200	.1668	.7944	.5366	.0338	.6875	.6538	.9819	.7422	.3619
.6300	.3276	.8619	.6353	.0395	.7090	.6695	.9923	.7734	.3623
.6400	.4895	.9296	.7344	.0453	.7305	.6853	3.0036	.8048	.3626
.6500	.6523	.9975	.8337	.0510	.7521	.7010	.0130	.8362	.3630
.6600	.8161	11.0656	.9334	.0568	.7736	.7168	.0234	.8678	.3634
.6700	.9810	.1339	16.0333	.0625	.7950	.7325	.0338	.8994	.3638
.6800	18.1468	.2024	.1336	.0683	.8165	.7483	.0441	.9312	.3641
.6900	.3137	.2711	.2341	.0740	.8380	.7640	.0545	.9630	.3645

TABLE 8-15. CONRADY G SUMS (*Continued*)

N	G_1	G_2	G_3	G_4	G_5	G_6	G_7	G_8	G_9
3.7000	18.4815	11.3400	16.3350	2.0797	6.8595	4.7797	3.0649	4.9950	0.3649
.7100	.6504	.4091	.4361	.0855	.8809	.7955	.0752	5.0270	.3652
.7200	.8202	.4784	.5376	.0912	.9024	.8112	.0856	.0592	.3656
.7300	.9911	.5479	.6393	.0969	.9238	.8269	.0960	.0914	.3660
.7400	19.1630	.6176	.7414	.1026	.9452	.8426	.1063	.1238	.3663
.7500	.3359	.6875	.8437	.1083	.9667	.8583	.1167	.1562	.3667
.7600	.5099	.7576	.9464	.1140	.9881	.8740	.1270	.1888	.3670
.7700	.6849	.8279	17.0493	.1197	7.0095	.8897	.1374	.2214	.3674
.7800	.8609	.8984	.1526	.1254	.0309	.9054	.1477	.2542	.3677
.7900	20.0379	.9691	.2561	.1311	.0523	.9211	.1581	.2870	.3681
.8000	.2160	12.0400	.3600	.1368	.0737	.9368	.1684	.3200	.3684
.8100	.3951	.1111	.4641	.1425	.0951	.9525	.1788	.3530	.3688
.8200	.5753	.1824	.5686	.1482	.1164	.9682	.1891	.3862	.3691
.8300	.7565	.2539	.6733	.1539	.1378	.9839	.1995	.4194	.3695
.8400	.9388	.3256	.7784	.1596	.1592	.9996	.2098	.4528	.3698
.8500	21.1221	.3975	.8837	.1653	.1805	5.0153	.2201	.4862	.3701
.8600	.3064	.4696	.9894	.1709	.2019	.0309	.2305	.5198	.3705
.8700	.4919	.5419	18.0953	.1766	.2232	.0466	.2408	.5534	.3708
.8800	.6783	.6144	.2016	.1823	.2445	.0623	.2511	.5872	.3711
.8900	.8659	.6871	.3081	.1879	.2659	.0779	.2615	.6210	.3715
.9000	22.0545	.7600	.4150	.1936	.2872	.0936	.2718	.6550	.3718
.9100	.2442	.8331	.5221	.1992	.3085	.1092	.2821	.6890	.3721
.9200	.4349	.9064	.6296	.2049	.3298	.1249	.2924	.7232	.3724
.9300	.6268	.9799	.7373	.2105	.3511	.1405	.3028	.7574	.3728
.9400	.8197	13.0536	.8454	.2162	.3724	.1562	.3131	.7918	.3731
.9500	23.0137	.1275	.9537	.2218	.3937	.1718	.3234	.8262	.3734
.9600	.2088	.2016	19.0624	.2275	.4149	.1875	.3337	.8608	.3737
.9700	.4049	.2759	.1713	.2331	.4362	.2031	.3441	.8954	.3741
.9800	.6022	.3504	.2806	.2387	.4575	.2187	.3544	.9302	.3744
.9900	.8005	.4251	.3901	.2444	.4787	.2344	.3647	.9650	.3747
4.0000	24.0000	.5000	.5000	.2500	.5000	.2500	.3750	6.0000	.3750
.0100	.2005	.5751	.6101	.2556	.5212	.2656	.3853	.0350	.3753
.0200	.4022	.6504	.7206	.2612	.5425	.2812	.3956	.0702	.3756
.0300	.6050	.7259	.8313	.2669	.5637	.2969	.4059	.1054	.3759
.0400	.8088	.8016	.9424	.2725	.5850	.3125	.4162	.1408	.3762
.0500	25.0138	.8775	20.0537	.2781	.6062	.3281	.4265	.1762	.3765
.0600	.2199	.9536	.1654	.2837	.6274	.3437	.4368	.2118	.3768
.0700	.4271	14.0299	.2773	.2893	.6486	.3593	.4471	.2474	.3771
.0800	.6355	.1064	.3896	.2949	.6698	.3749	.4575	.2832	.3775
.0900	.8449	.1831	.5021	.3005	.6910	.3905	.4678	.3190	.3778
.1000	26.0555	.2600	.6150	.3061	.7122	.4061	.4780	.3550	.3780
.1100	.2672	.3371	.7281	.3117	.7334	.4217	.4883	.3910	.3783
.1200	.4801	.4144	.8416	.3173	.7546	.4373	.4986	.4272	.3786
.1300	.6940	.4919	.9553	.3229	.7757	.4529	.5089	.4634	.3789
.1400	.9092	.5696	21.0694	.3285	.7969	.4685	.5192	.4998	.3792
.1500	27.1254	.6475	.1837	.3340	.8181	.4840	.5295	.5362	.3795
.1600	.3428	.7256	.2984	.3396	.8392	.4996	.5398	.5728	.3798
.1700	.5614	.8039	.4133	.3452	.8604	.5152	.5501	.6094	.3801
.1800	.7811	.8824	.5286	.3508	.8815	.5308	.5604	.6462	.3804
.1900	28.0020	.9611	.6441	.3563	.9027	.5463	.5707	.6830	.3807
.2000	.2240	15.0400	.7600	.3619	.9238	.5619	.5810	.7200	.3810

8.5. Useful Equations

8.5.1. Reflection and Transmission at a Single Surface. E_{np}^{+} is the electric vector of a wave [7] traveling in the positive direction in the nth layer polarized with the electric vector parallel to the plane of incidence: E_{ns}^{-} is the electric vector in the negative direction in the nth layer polarized perpendicular to the plane of incidence. Let n be $n - ik$. Then the Fresnel equations are:

$$\frac{E_{0p}^{-}}{E_{0p}^{+}} = \frac{n_0 \cos \phi_1 - n_1 \cos \phi_0}{n_0 \cos \phi_1 + n_1 \cos \phi_0} = r_{1p} \tag{8-3}$$

$$\frac{E_{1p}^{+}}{E_{0p}^{+}} = \frac{2n_0 \cos \phi_0}{n_0 \cos \phi_1 + n_1 \cos \phi_0} = t_{1p} \tag{8-4}$$

$$\frac{E_{0s}^{-}}{E_{0s}^{+}} = \frac{n_0 \cos \phi_0 - n_1 \cos \phi_1}{n_0 \cos \phi_0 + n_1 \cos \phi_1} = r_{1s} \tag{8-5}$$

$$\frac{E_{1s}^{+}}{E_{0s}^{+}} = \frac{2n_0 \cos \phi_0}{n_0 \cos \phi_0 + n_1 \cos \phi_1} = t_{1s} \tag{8-6}$$

Sometimes a more useful form for amplitude relations is

$$r_{1p} = \frac{\tan (\phi_1 - \phi_0)}{\tan (\phi_1 + \phi_0)} \tag{8-7}$$

$$t_{1p} = \frac{2 \sin \phi_1 \cos \phi_0}{\sin (\phi_1 + \phi_0) \cos (\phi_1 - \phi_0)} \tag{8-8}$$

$$r_{1s} = \frac{\sin (\phi_1 - \phi_0)}{\sin (\phi_1 + \phi_0)} \tag{8-9}$$

$$t_{1s} = \frac{2 \sin \phi_1 \cos \phi_0}{\sin (\phi_1 + \phi_0)} \tag{8-10}$$

For intensity,

and

$$\left.\begin{aligned} \rho_p &= \frac{(E_{0p}^{-})^2}{(E_{0p}^{+})^2} = r_{1p}^2 \\ \rho_s &= \frac{(E_{0s}^{-})^2}{(E_{0s}^{+})^2} = r_{1s}^2 \end{aligned}\right\} \tag{8-11}$$

and the transmittances are given by

and

$$\left.\begin{aligned} \tau_p &= \frac{n_1 (E_{1p}^{+})^2}{n_0 (E_{0p}^{+})^2} = \frac{n_1}{n_0} t_{1p}^2 \\ \tau_s &= \frac{n_1 (E_{1s}^{+})^2}{n_0 (E_{0s}^{+})^2} = \frac{n_1}{n_0} t_{1s}^2 \end{aligned}\right\} \tag{8-12}$$

Then for normal incidence,

$$\rho_p = \rho_s = \left(\frac{n_0 - n_1}{n_0 + n_1}\right)^2$$

$$\tau_p = \tau_s = \frac{4\,n_0 n_1}{(n_0 + n_1)^2} \tag{8-13}$$

For absorbing layers the equations are complicated. For normal incidence,

$$r_{1p} = r_{1s} = \frac{n_0 - n_1 + ik_1}{n_0 + n_1 - ik_1} \tag{8-14}$$

which gives, for the reflectance of the surface

$$\rho_p = \rho_s = \frac{(n_0 - n_1)^2 + k_1^2}{(n_0 + n_1)^2 + k_1^2} \tag{8-15}$$

When $n^2 + k^2 >> 1$,

$$\rho_p = \frac{(n^2 + k^2)\cos^2\phi_0 - 2n\cos\phi_0 + 1}{(n^2 + k^2)\cos^2\phi_0 + 2n\cos\phi_0 + 1} \tag{8-16}$$

$$\rho_s = \frac{(n^2 + k^2) - 2n\cos\phi_0 + \cos^2\phi_0}{(n^2 + k^2) + 2n\cos\phi_0 + \cos^2\phi_0} \tag{8-17}$$

8.5.2. Reflection and Transmission by a Single Layer. When multiple traversals are taken into account, the amplitude reflection coefficient is

$$r = r_1 + t_1 t'_1 r_2 e^{-2i\delta_1} - t_1 t'_1 r_1 r_2^2 e^{-4i\delta_1} + \ldots$$

or

$$= r_1 + \frac{t_1 t'_1 r_2 e^{-2i\delta_1}}{1 + r_1 r_2 e^{-2i\delta_1}} \tag{8-18}$$

The amplitude transmission coefficient is

$$t = t_1 t_2 e^{-i\delta_1} - t_1 t_2 r_1 r_2 e^{-3i\delta_1} + t_1 t_2 r_1^2 r_2^2 e^{-5i\delta_1} - \ldots$$

$$= \frac{t_1 t_2 e^{-i\delta_1}}{1 + r_1 r_2 e^{-2i\delta_1}} \tag{8-19}$$

To account for polarization differences, expressions for t_1, t_2, r_1, and r_2 should be obtained from the proper equations (8-3 to 8-6). The intensity ratios are

$$\rho = \frac{r_1^2 + 2r_1 r_2 \cos 2\delta_1 + r_2^2}{1 + 2r_1 r_2 \cos 2\delta_1 + r_1^2 r_2^2} \tag{8-20}$$

$$\tau = \frac{n_2}{n_0} \cdot \frac{t_1^2 t_2^2}{(1 + 2r_1 r_2 \cos 2\delta_1 + r_1^2 r_2^2)} \tag{8-21}$$

where δ_1 is $(2\pi/\lambda)n_1 d_1 \cos\phi_1$. For normal incidence,

$$\rho = \frac{(n_0^2 + n_1^2)(n_1^2 + n_2^2) - 4n_0 n_1^2 n_2 + (n_0^2 - n_1^2)(n_1^2 - n_2^2)\cos 2\delta_1}{(n_0^2 + n_1^2)(n_1^2 + n_2^2) + 4n_0 n_1^2 n_2 + (n_0^2 - n_1^2)(n_1^2 - n_2^2)\cos 2\delta_1} \tag{8-22}$$

$$\tau = \frac{8n_0 n_1^2 n_2}{(n_0^2 + n_1^2)(n_1^2 + n_2^2) + 4n_0 n_1^2 n_2 + (\mathrm{n}_0^2 - n_1^2)(n_1^2 - n_2^2)\cos 2\delta_1} \tag{8-23}$$

For absorbing media n must be replaced by $\tilde{n}$.

8.5.3. Emissivity, Transmissivity, and Reflectivity for Partially Transparent Bodies [8]. The emissivity, ϵ, apparent (measured) reflectance, ρ_m, and apparent (measured) transmittance, τ_m, of partially transparent bodies can be related as follows [8]:

$$\epsilon + \tau_m + \rho_m = 1 \tag{8-24}$$

These terms are given below, where the functional dependence of all quantities on temperature and wavelength is suppressed.

$$\epsilon = \frac{(1 - \rho)(1 - \tau)}{1 - \rho\tau} \tag{8-25}$$

$$\rho_m = \rho + \frac{\rho\tau^2 (1 - \rho)^2}{1 - \rho^2\tau^2} \tag{8-26}$$

$$\tau_m = \tau \frac{(1 - \rho)^2}{1 - \rho^2\tau^2} \tag{8-27}$$

Figure 8-19 is a plot of Eq. (8-24). The complete thermal radiation properties of a given body at a particular wavelength and for a particular temperature are represented by a single point on this chart. The location of the point can be determined experimentally when any two of the three quantities ϵ, τ_m, and ρ_m are given. From this, values of τ and ρ can be obtained.

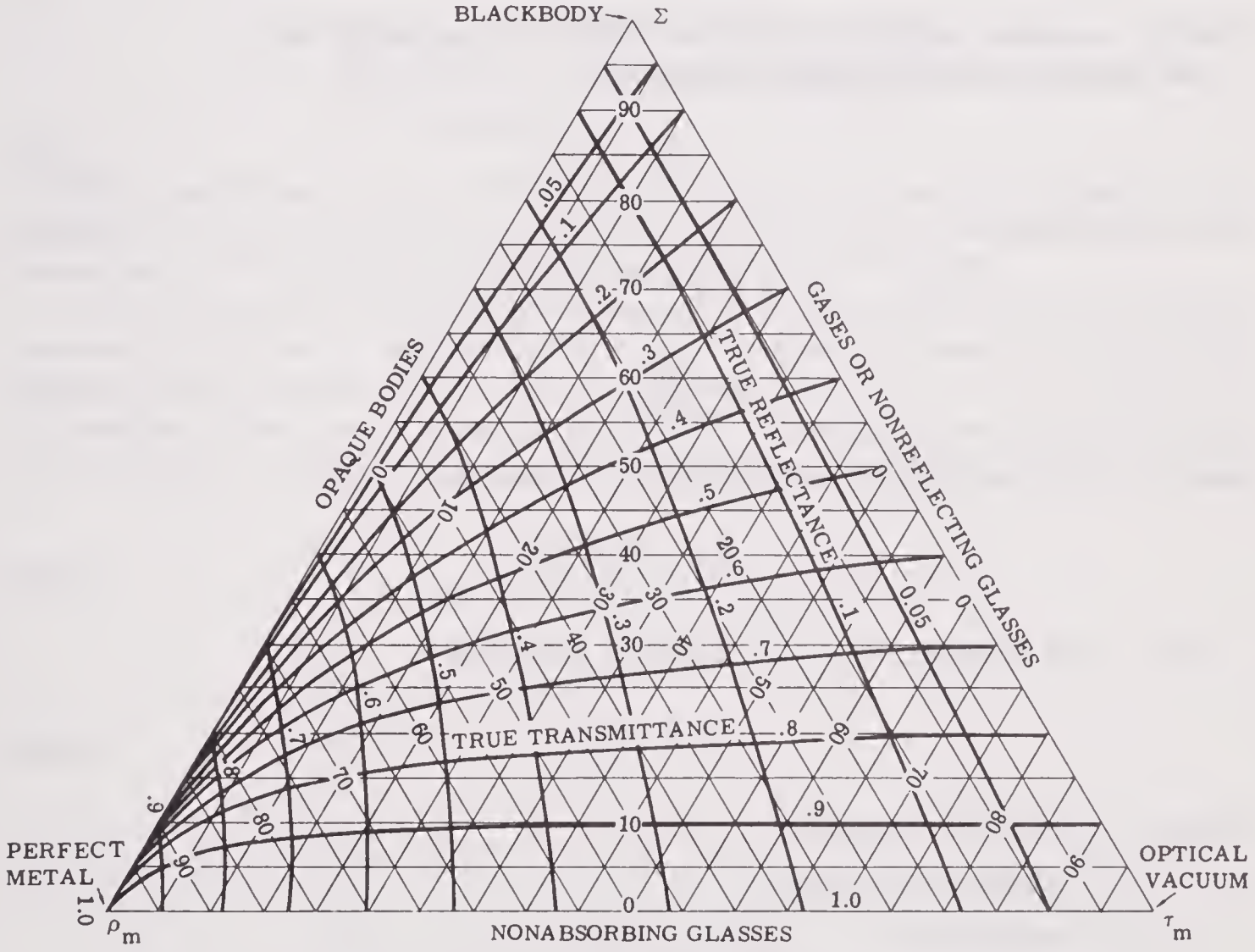


FIG. 8-19. Thermal radiation chart for determining true values of transmissivity τ and reflectivity ρ from experimentally measured values of emissivity ϵ, apparent transmissivity τ_m, and apparent reflectivity ρ_m.

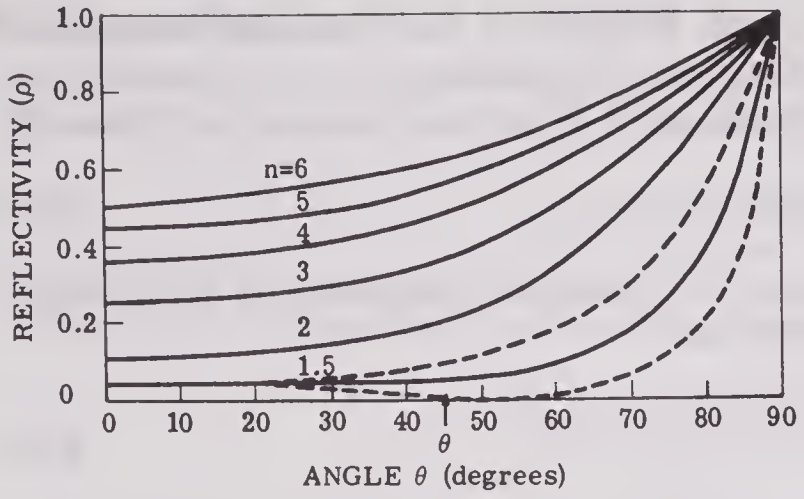


FIG. 8-20. Reflection loss for different incidence angles and different refractive indices.

Figure 8-20 illustrates Fresnel's laws of reflection. It shows reflectivity ρ as a function of angle for different refractive indices. The lower dashed curve shows ρ_p, and the upper dashed curve ρ_s, where subscripts p and s indicate the parallel and the perpendicular components of the electric field, respectively. The curve between them is characteristic of unpolarized light at the angles indicated for $n = 1.5$. The other curves are also for unpolarized light. The equations and curves show that there is a phase change of π for E_s vibrations of all angles of incidence. There is a phase change for E_p vibrations only for angles equal to or exceeding the polarizing angle. For normal incident radiation the equation for reflectivity reduces to

$$\rho = \left(\frac{n-1}{n+1}\right)^2 \tag{8-28}$$

Finally, at grazing incidence angles, the reflectivity approaches 100%.

The transmittance for multiple reflections is

$$\tau_m = \frac{(1-\rho)^2 e^{-ax}}{1-\rho^2 e^{-2ax}} \tag{8-29}$$

If a is small, then

$$\tau_\infty = \frac{(1-\rho)^2}{1-\rho^2} = \frac{1-\rho}{1+\rho}$$

$$= \frac{2n}{n^2+1} \tag{8-30}$$

where τ_∞ = transmission for infinite number of reflections. Also,

$$\rho_\infty = \frac{(n-1)^2}{n^2+1} \tag{8-31}$$

8.5.4. Loss Tangent. The loss tangent is defined as

$$\tan \delta = \frac{I_c}{I_d} = \frac{\sigma}{\epsilon\omega} \tag{8-32}$$

where I_c = conduction current

I_d = displacement current

σ = conductivity

ϵ = dielectric constant

ω = angular frequency

The loss tangent under proper assumptions is related to the absorption coefficient as follows:

$$\tan \delta = a\lambda/2\pi n$$

8.5.5. Extinction Coefficient. The extinction coefficient k is defined as

$$k = a\lambda/4\pi \tag{8-33}$$

8.6. Optical Surface Coatings

8.6.1. Reflective Coatings. Aluminum, silver, gold, copper, rhodium, and titanium are the most frequently used mirror metals. Figure 8-21 shows the measured reflectance of films of these metals.

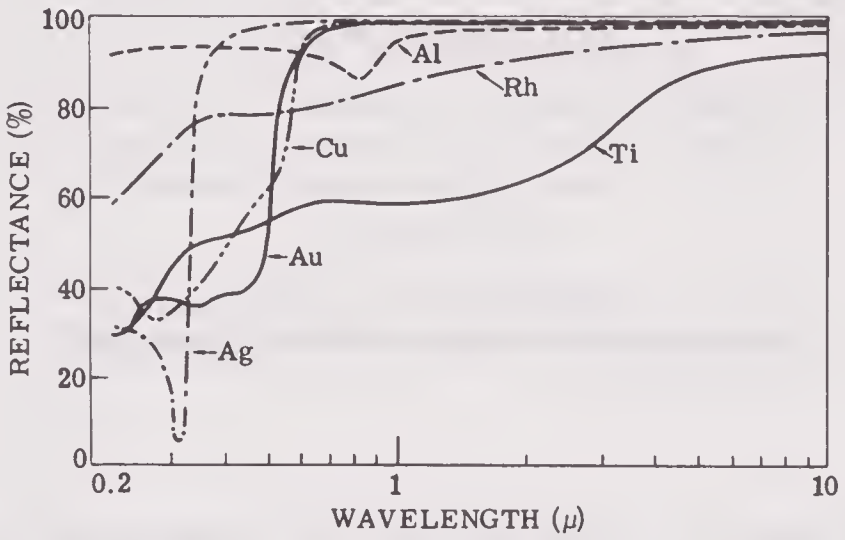


FIG. 8-21. Reflectance of various films of silver (Ag), gold (Au), aluminum (Al), copper (Cu), rhodium (Rh), and titanium (Ti) as a function of wavelength [9].

Stellite* cobalt-base alloy is often used when the surface must be exposed to the elements. It resists tarnish and corrosion and can be cleaned easily with a chalk-acetone paste. The reflectivity of Stellite* is given in Fig. 8-22 [10].

The most frequently used high-reflecting coating for first-surface mirrors is vacuum-deposited aluminum. It adheres better to glass and other substrates than the other high reflecting materials, does not tarnish in normal air, and is very easy to evaporate.

The effect of the speed of evaporation or deposition rate on the reflectance of aluminum coatings is considerable at visible and shorter wavelengths; however, in the

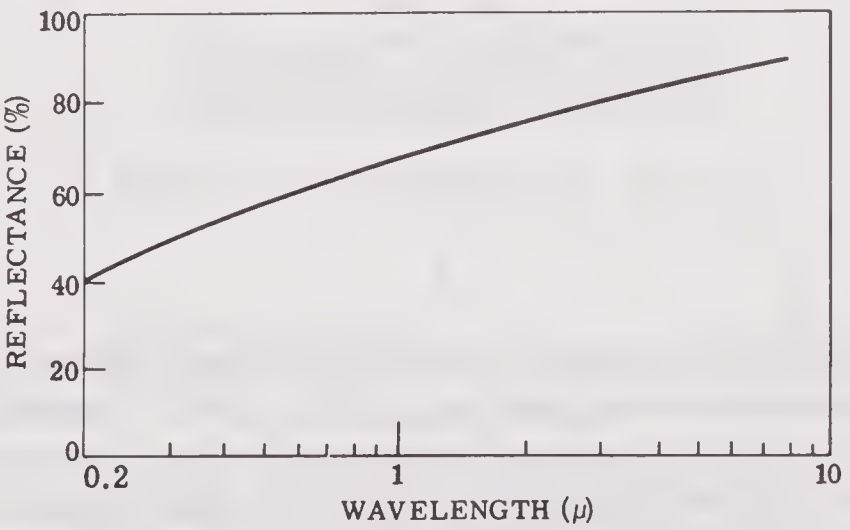


FIG. 8-22. Reflectance of Stellite* [10].

*Registered trademark of Union Carbide Corporation.

infrared this becomes much less pronounced (Fig. 8-23). Almost any evaporation speed within reasonable limits results in aluminum coatings with more than 95% reflectance between 2 and 10 μ. It is always advisable, however, to evaporate aluminum as fast as possible because fast-deposited films are chemically and mechanically more durable.

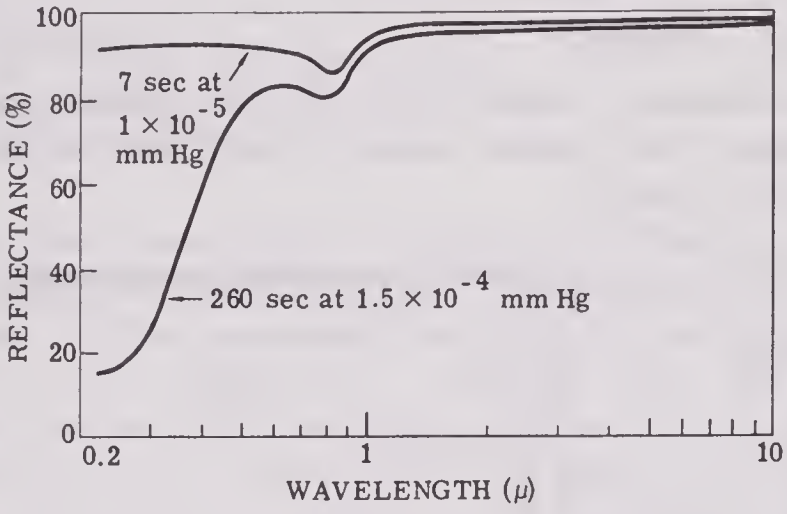


FIG. 8-23. Reflectance of two aluminum films prepared under extremely different evaporation conditions [9].

8.6.2. Filter Mirrors [9]. There are two film combinations that provide high infrared but low visible reflectance. These are:

(1) Aluminum coated with evaporated germanium monoxide and silicon monoxide, each film approximately one quarter wavelength thick (Fig. 8-24). When germanium is evaporated onto opaque aluminum, the aluminum reflectance, controlled at a certain wavelength in the visible, decreases to a minimum of 30% to 40%. The addition of silicon monoxide decreases the visible reflectance to almost zero. Figure 8-25 shows the reflectance of two such mirrors.

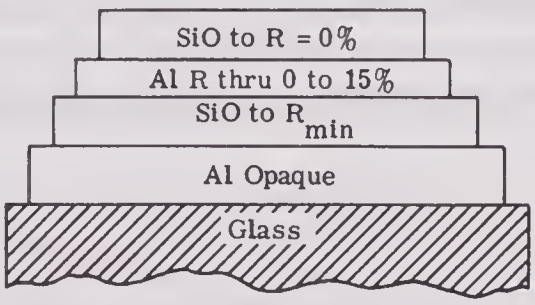


FIG. 8-24. Layer arrangement in $(Al\text{-}SiO)^2$ filter mirror [9].

(2) Two silicon monoxide coatings, separated by a semitransparent aluminum film on top of the opaque aluminum mirror coating (Fig. 8-26). The first silicon monoxide film is evaporated until the aluminum reflectance reaches a minimum of about 60%. Then the aluminum is deposited until the reflectance, after passing through zero, rises again to a value of 15% to 20%. Another film of silicon monoxide on top brings the reflectance again down to zero. Figure 8-27 shows the reflectance of this type of reflector in the visible and infrared.

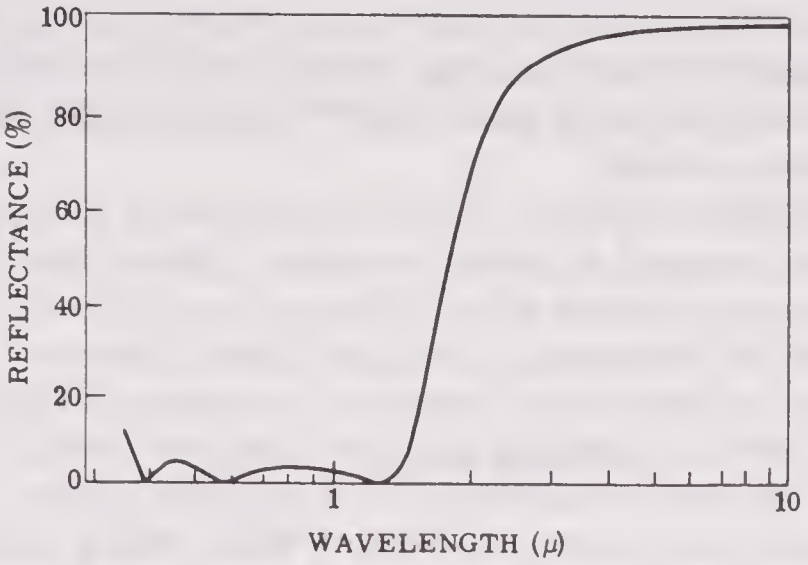


FIG. 8-25. Reflectance of Al-SiO filter mirror as a function of wavelength. Layer deposition controlled 0.55 μ [9].

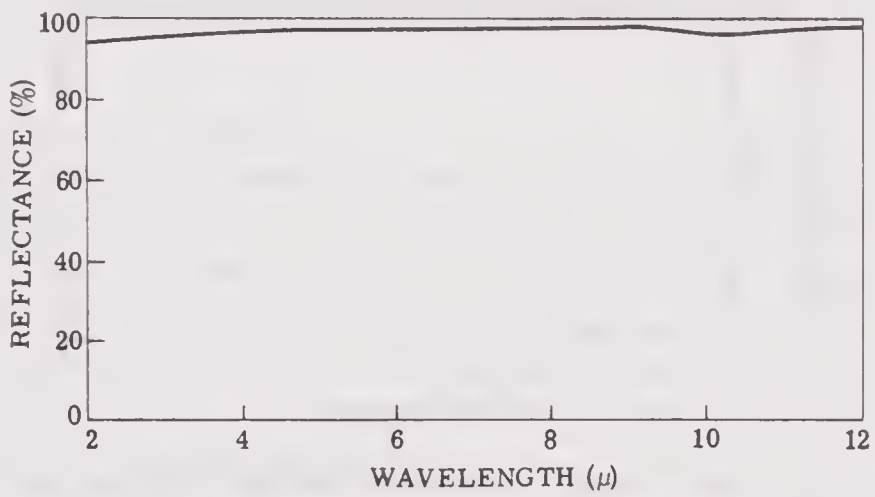


FIG. 8-26. Layer arrangement in Al-SiO-Al-SiO filter mirror [9].

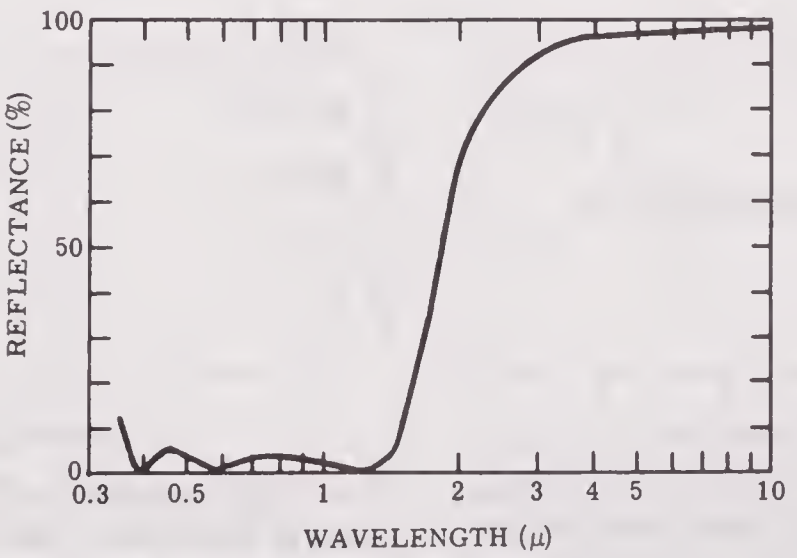


FIG. 8-27. Reflectance of Al-SiO-Al-SiO filter mirror as a function of wavelength. Layer deposition controlled at 550 mμ [9].

8.6.3. Protective Coatings [9]. For many mirrors, the natural oxide film that forms on an evaporated aluminum surface is too thin to furnish sufficient protection, especially if the mirror requires frequent cleaning. Silicon monoxide affords a suitable hard protective coat. It evaporates at about 1200°C and condenses on the mirror surface in uniform and adherent layers.

There is a considerable difference in the characteristics of coatings when used for reflective optics as compared to refractive optics. Figure 8-28 shows the infrared reflectance of aluminum protected with a 125-mm film of SiO.

Figure 8-29 shows the reflectance curve of an "Alzac" reflector. The Alzac reflector is an aluminum mirror form whose surface is electrolytically polished or brightened and then protected with an anodically produced aluminum oxide film about 1.4 μ thick. Figure 8-30 shows the effect of applying a 1.2-μ-thick film of SiO. At 11 μ, where the SiO film is effectively one quarter wavelength thick, strong absorption takes place and mirror reflectance is decreased to almost zero.

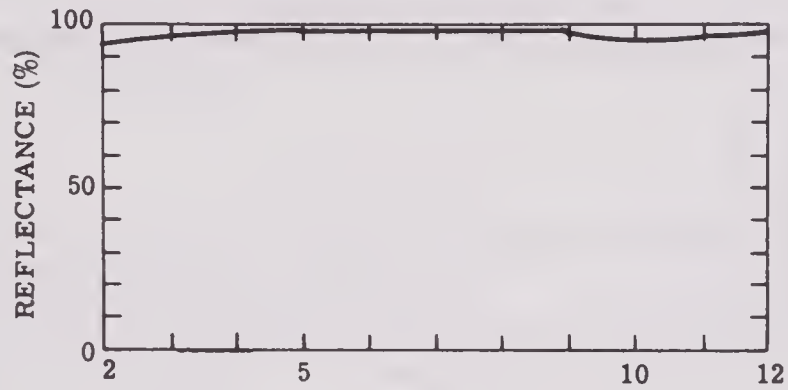


FIG. 8-28. Infrared reflectance of aluminum protected with a 125-mμ film of silicon oxide [9].

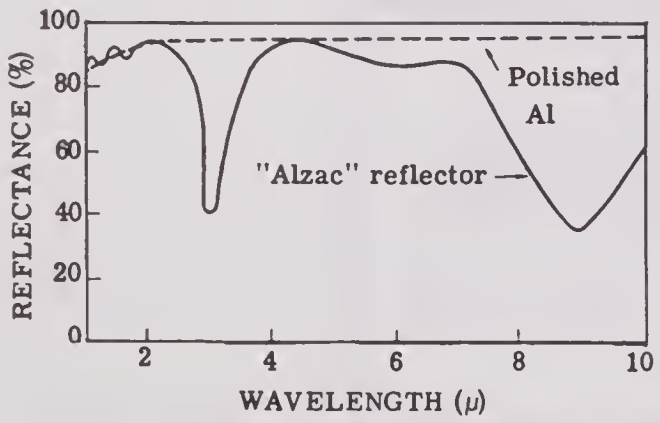


FIG. 8-29. Infrared reflectance of polished aluminum and of an "Alzac" reflector (aluminum electrolytically brightened and protected with about 1.4 μ of Al_2O_3) [9].

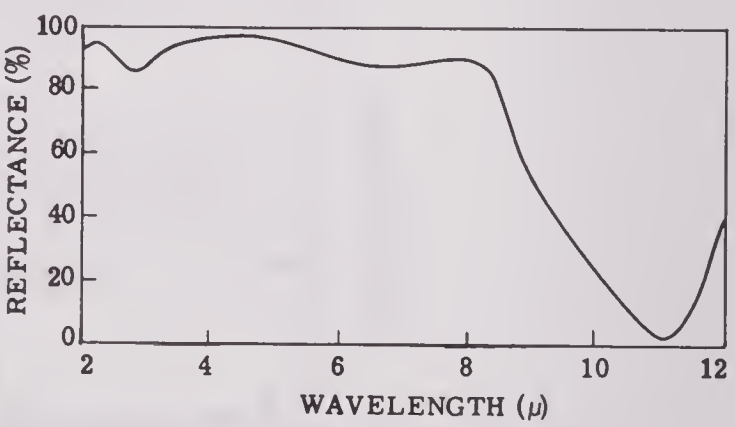


FIG. 8-30. Reflectance of aluminum protected with a 1.2-μ-thick film of silicon oxide (SiO deposited in 20 min at 1×10^{-5} mm Hg) [9].

8.6.4. Antireflection Coatings [7,9]. Infrared antireflection coatings increase the energy transmitted through optical surfaces by reducing Fresnel reflection losses. The criteria for an antireflection coating are (1) that its refractive index be equal to the square root of the substrate index when the substrate is in air:

$$n = \sqrt{n'} \tag{8-34}$$

and (2) that the phase difference between the incident wave and the reflected wave be an odd multiple of π:

$$\Delta\phi = 4\pi nd \cos r/\lambda = 2\pi\,(m + \tfrac{1}{2})$$

which reduces to

$$nd \cos r = (\lambda/4)\,(2m + 1) \tag{8-35}$$

where n = refractive index of film
n' = refractive index of material
$\Delta\phi$ = phase change of incident light
λ = wavelength of incident light
d = thickness of film
r = angle of inclination in the film of the ray to the film normal
m = 1, 2, 3, . . .

For more details, see Chapter 7.

Table 8-16 lists typical antireflection coatings and their effect on the infrared transmission of various materials. The transmittance curves of silicon (Fig. 8-31) and germanium (Fig. 8-32) are shown with and without coatings. Figure 8-33 shows the measured transmittance curve of an arsenic trisulfide plate coated on both sides with double quarter-wave films. Figure 8-34 shows the measured transmittance and reflectance of a glass plate coated on each side with a half-wave film of silicon oxide and a quarter-wave film of magnesium fluoride.

TABLE 8-16. EXAMPLES OF ANTIREFLECTION COATING [11]

Sample and Thickness (in.)		*Uncoated Transmission*		*Coated Transmission*		
				MgF_2	LiF	
		at 2 μ	*at* 3.6 μ	*at* 2 μ	*at* 2 μ	*at* 3.6 μ
Fuzed quartz	0.203	0.935		0.97	0.97	
Dense flint	0.061	0.855		0.955		
MgO	0.129	0.86	0.85	0.945		0.94
Al_2O_3	0.063	0.88		0.985		
Al_2O_3	0.126		0.875			0.97

8.6.5. Replica Mirrors [9]. A simple and inexpensive method for producing replicas of optical elements, such as mirrors and trihedral prisms, with plastic backings is described in reference [9]. In this technique replicas equipped with the final high-reflecting coatings are prepared directly on the master mold. Epoxy resins are used to make the replica mirror form. The epoxy casting compounds are pourable liquids that can be hardened overnight with little or no heat through the addition of a liquid hardener. Very little heat and very low shrinkage are developed during the hardening process.

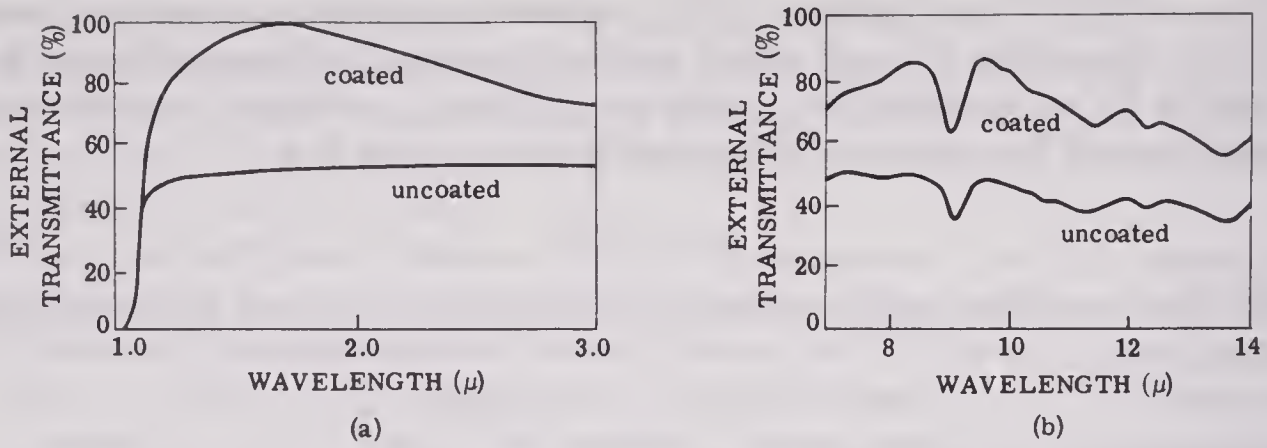


FIG. 8-31. Transmittance of a silicon plate, 1.5 mm thick with antireflection coatings of: (*a*) SiO ($n = \lambda/4$ at 1.7 μ); (*b*) ZnS ($n = \lambda/4$ at 9.8 μ) [12].

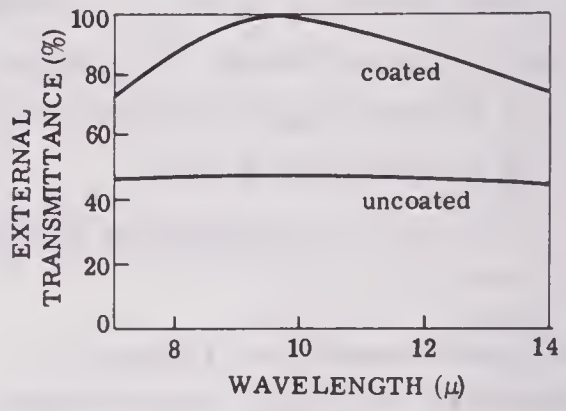


FIG. 8-32. Transmittance of a germanium plate, 1.0 mm thick, with antireflection coating of ZnS ($n = \lambda/4$ at 9.8 μ).

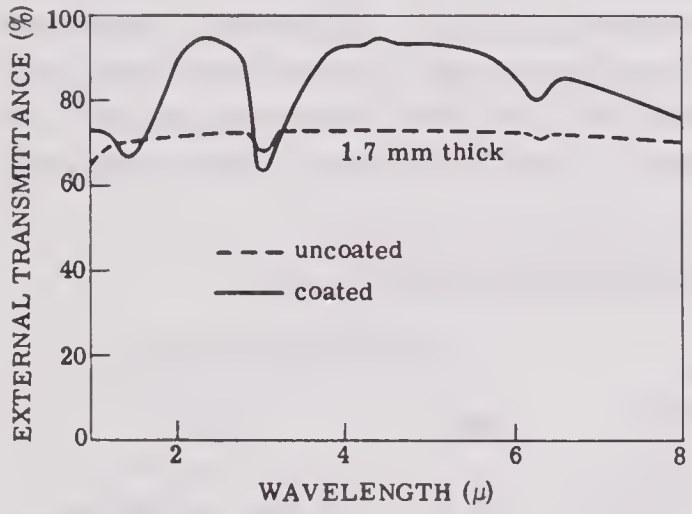


FIG. 8-33. Transmittance of an As_2S_3 plate ($n = 2.4$) antireflected with double $\lambda/4$ films of tungsten oxide ($n = 1.8$) and sodium aluminum fluoride ($n = 1.35$), each $\lambda/4$ thick at 3 μ [9].

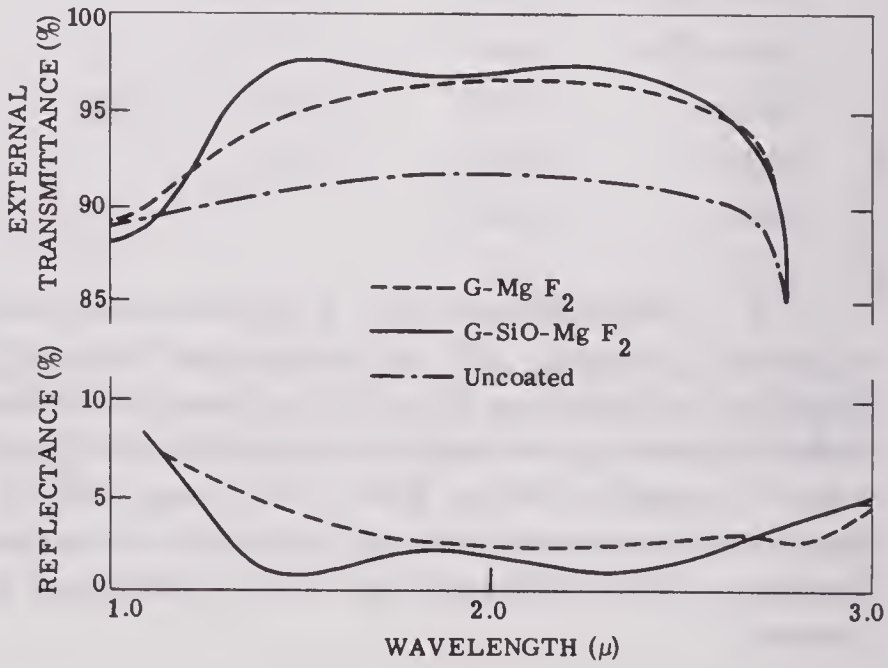


FIG. 8-34. Transmittance and reflectance of a glass plate with a $\lambda/2 - \lambda/4$ SiO ($n = 1.7$) and MgF_2 ($n = 1.38$) coating, in comparison with a $\lambda/4$ MgF_2 coating [9].

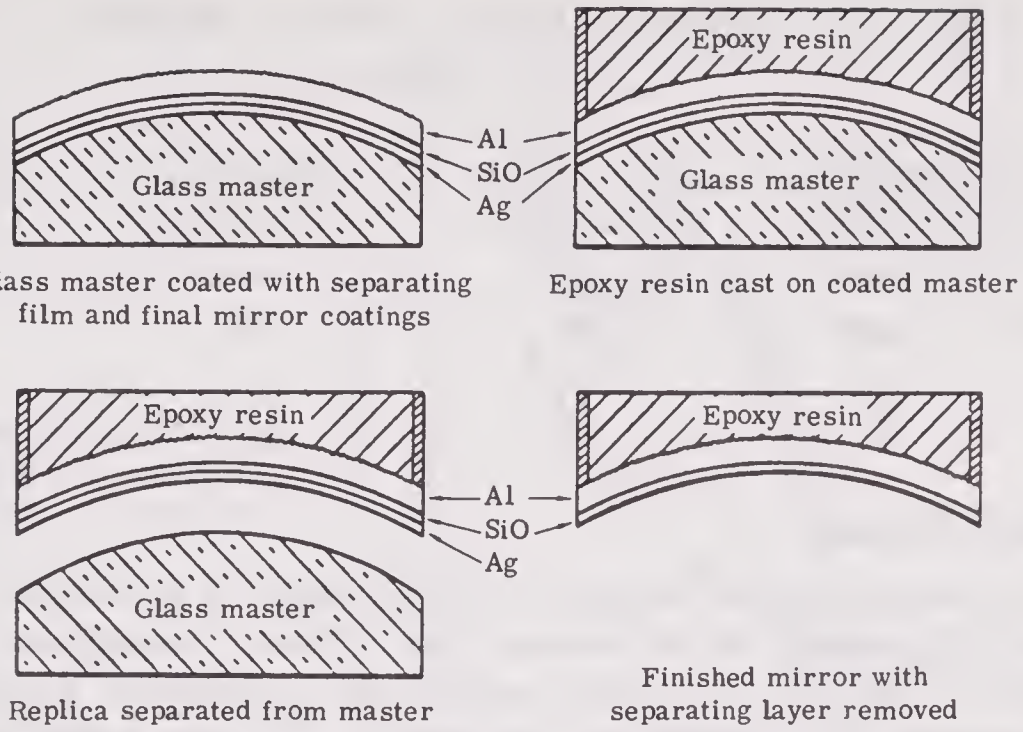


FIG. 8-35. The four steps for preparing protected replica mirrors with epoxy backing [9].

The steps involved in producing replica mirrors with epoxy resins are shown in Fig. 8-35.

In making a replica mirror, the negative glass form is coated with silver, silicon monoxide, and a heavy deposit of aluminum. A plastic mold in the form of a ring is then placed on the coated glass master and filled with the epoxy resin. After the epoxy resin has cured, the replica is separated from the master by applying slight pressure to the plastic ring. The silver film, which has poor adherence to glass, is used only to make the separation easy. After the separation, the silver film is removed in nitric acid, leaving the replica mirror coated with aluminum protected with silicon monoxide. The aluminum, which has been deposited rather heavily to produce a rough surface, adheres excellently to the hardened epoxy compound. The quality of replica optics made with epoxy resins can be greatly improved by adding equal or even higher amounts of finely divided silica to the epoxy-type molding compound. The addition of silica fillers reduces shrinkage during the hardening process and results in mirror forms with increased hardness and heat resistance, and reduced expansion coefficients. On an optical test bench, the replica mirror can be compared with the original mirror by focusing incident collimated light on pinhole apertures of various sizes and measuring the ratio of the light transmitted by the apertures to the total amount of light in the focal plane. The performance of paraboloidal plastic replicas as compared with the original glass mirrors is given in Table 8-17. The effects of aging on replica mirrors is given in Table 8-18.

TABLE 8-17. PERFORMANCE OF PLASTIC REPLICA MIRRORS [13]

Aperture Size (mm)	*Energy through Aperture* (%)		
	Original Glass Mirror	*Replica from Master No. 1*	*Replica from Master No. 2*
0.79	98	96	97
0.38	96	89	90
0.27	82	81	80
0.18	60	59	61

TABLE 8-18. EFFECTS OF AGING ON REPLICA MIRRORS [13]

Aperture Size (mm)	*Energy through Aperture* (%)			
	Original Glass Mirror	*Replica No. 1* (2 *days old*)	*Replica No. 2* (1 *day old*)	*Replica No. 2* (19 *months old*)
0.79	100	99	98	98
0.38	99	97	98	98
0.27	99	97	96	95
0.18	97	95	94	88

8.7. Radiation Damage

Protons and electrons inflict damage at a certain depth in a material. This depth is a function of the density of the material itself. Thus an absorption cross section which has the units g^{-1} cm^3 is a useful concept, and a penetration "depth" in units of g cm^{-2} is a measure of the susceptibility to damage. Whereas protons and electrons damage at a given depth, the number of photons and the resultant damage decreases exponentially with distance. Thus photon range is defined as the g cm^{-2} at which the flux is reduced by a factor of e. Figure 8-36 presents the resistance of various materials to radiation damage, and Fig. 8-37 is a plot of silicon absorption coefficient as affected by neutron bombardment.

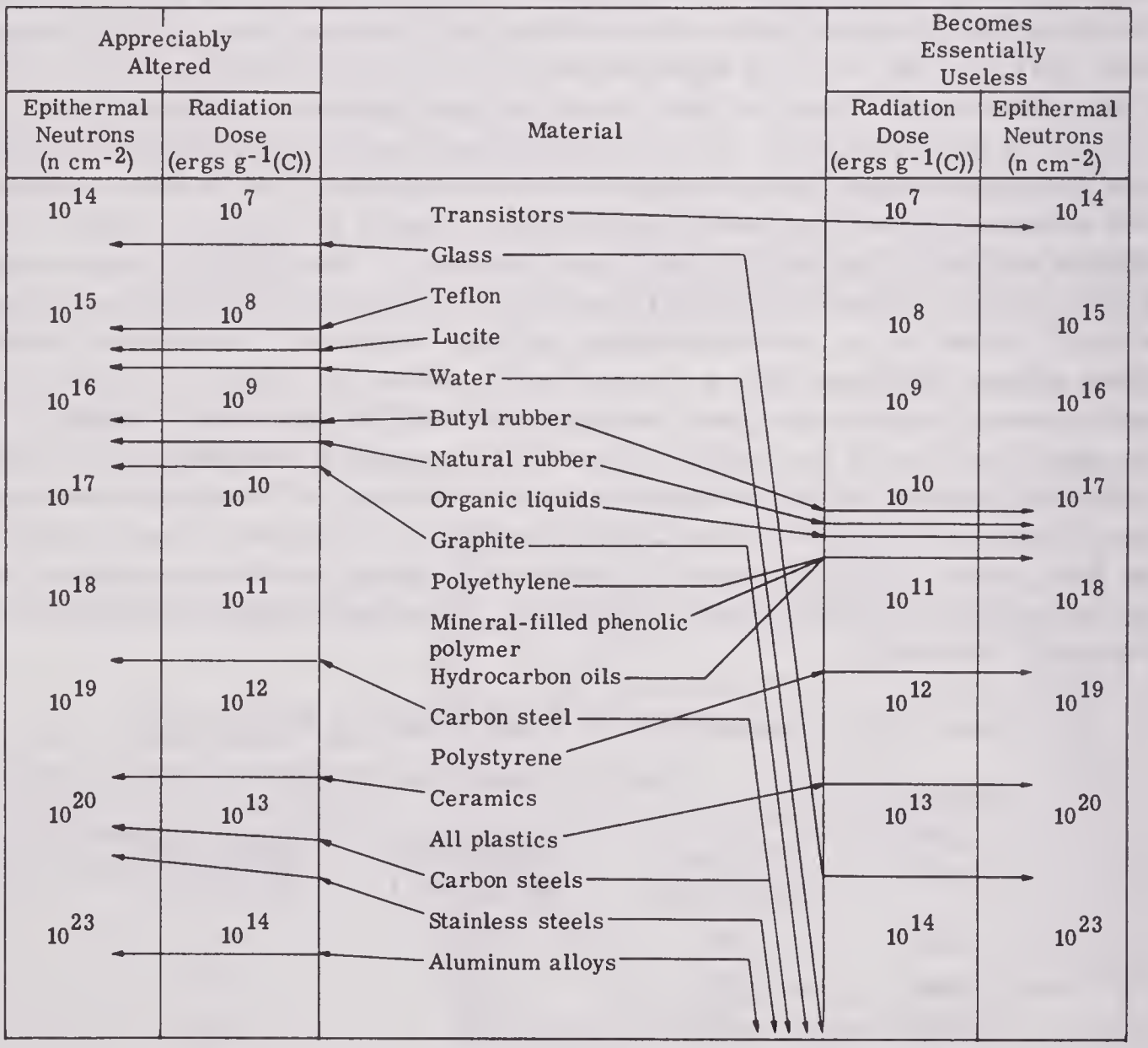


FIG. 8-36. Radiation resistance of various types of materials [14].

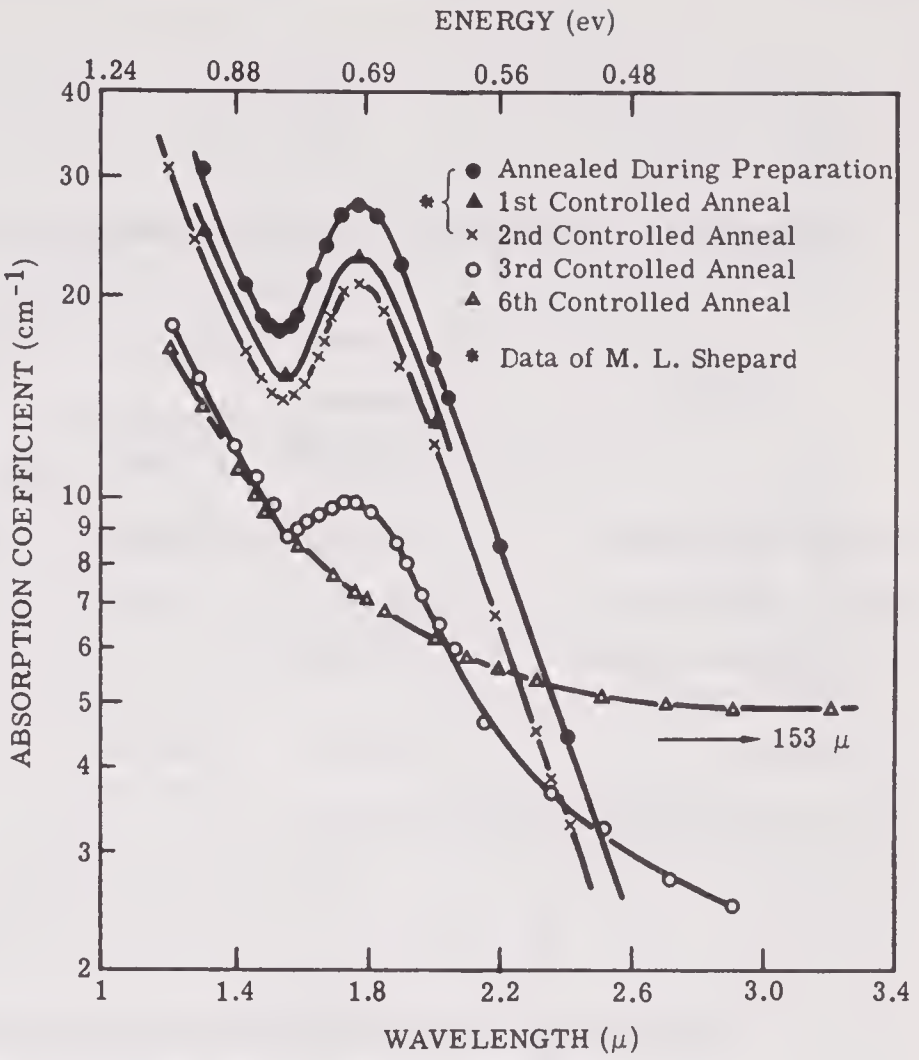


FIG. 8-37. Absorption coefficient *vs.* wavelength for neutron-bombarded silicon at room temperature [15].

Radiation causes ionization and atomic displacement. The rate of ionization can be given in erg g^{-1} yr^{-1}, and displacement in fractions of atoms displaced per year. Tables 8-19 and 8-20 [16] provide general data on the performance of materials. More specific data can be found in other references [17 to 70].

8.8. Infrared Transmission of Cooled Optical Materials [71]

Figure 8-38 shows the transmittance of a number of cooled optical materials in the 1-5.5-μ region. In most cases, the cooled samples transmit more than the uncooled samples. If the increase in transmission due to the decrease of Fresnel reflection loss at the sample surface is subtracted from the measured total increase of transmission, the absolute increase in transmission of the material is obtained. Some of the increased transmission results from the reduced path length through the coolant, a function of the sample thickness. However, calculations [71] show that this increase can never be greater than 0.8% for the thickest sample.

8.9. Optical Properties of Blacks

Spectral emittances of blacks have been measured by Stierwalt, Kirk, and Bernstein [72]. The data are plotted in Fig. 8-39. Additional data are given by Harris and Cuff [73] on specular and diffuse reflectances – including goniophotometric curves – for acetylene smoke black, Dupont flat black (unspecified), Eastman Kodak NOD-18, Goldblack, Lampblack, soot, and lacquer. All were deposited on glass microscope slides and measured from 0.254 to 1.1 μ.

Other data on the emissivity (emissance) of materials can be found in an Arthur D. Little report [74]. They include measurements of Globars, BN, SiC, zirconia, magnesia, alumina, aluminum, Inconel, stainless steel; also white enamel and black enamel on Inconel, and SiO on platinum and Inconel. The data are not for blacks directly, but they may be useful.

TABLE 8-19. RADIATION DOSAGES PRODUCED BY ATOMIC PARTICLES IN SPACE

Radiation	*Ionization* (erg g^{-1} yr^{-1})		*Fraction of Atoms Displaced per Year*		
	Surface and Through 1 mg cm^{-2}	*Through* 1 g cm^{-2}	*Exposed Surface*	*Through* 1 mg cm^{-2}	*Through* 1 g cm^{-2}
Inner radiation belt	10^{14}(?)	10^{7}-10^{8}	10^{-1}(?)	10^{-5}	10^{-9}
Outer radiation belt	10^{13}-10^{15}	10^{6}-10^{8}	10^{-12}-10^{-10}	10^{-12}-10^{-10}	10^{-13}
Solar emission (flares except as noted)	10^{7}-10^{9}(?)	10^{4}-10^{5}	10^{-12}-10^{0}[a]	10^{-12}-10^{-11}	10^{-13}-10^{-11}
Cosmic rays	10^{2}-10^{3}	10^{2}-10^{3}	10^{-14}-10^{-13}	10^{-14}-10^{-13}	10^{-14}-10^{-13}

[a]May be displaced by steady solar emission.

TABLE 8-20. RADIATION DOSE PRODUCING APPRECIABLE CHANGE IN ENGINEERING PROPERTIES OF VARIOUS MATERIALS

Material	*Changed Properties*	*Ionization* (erg g^{-1})	*Fraction of Atoms Displaced*
Plastics			
Tetrafluoroethylene	In air, mechanical, electrical	10^{6}-10^{7}	–
	No air, mechanical, electrical	10^{7}-10^{9}	–
Other	Optical transmission	10^{6}-10^{11}	–
	Dimensions, mechanical, electrical[a]	10^{7}-10^{11}	–
Elastomers	Mechanical	10^{8}-10^{10}	–
Oils and greases	Lubrication, consistency	10^{9}-10^{12}	–
Ceramics			
Glass	Optical transmission	10^{5}-10^{10}	10^{-11}-10^{-7}
	Dimensions, mechanical	$\geq 10^{11}$	$\sim 10^{-7}$
	Electrical	$>10^{11}$	10^{-7}-10^{-6}
Fused silica	Optical transmission	10^{7}-10^{11}	10^{-9}-10^{-5}
Crystalline	Optical transmission	10^{5}-10^{11}	10^{-8}-10^{-4}
	Dimensional, mechanical	$>10^{11}$	10^{-4}-10^{-2}
	Electrical[b]	$>10^{11}$	10^{-3}-10^{-1}
Semiconductor (devices)	Minority carrier effects	–	10^{-12}-10^{-10}
	Majority carrier effects	–	10^{-9}-10^{-6}
Metals	Ferromagnetism	–	10^{-6}-10^{-5}
	Mechanical	–	10^{-4}-10^{-3}
	Electrical	–	10^{-3}-10^{-2}

[a]Temporary increase in electrical conductivity during irradiation at dose rates $\sim 10^{6}$ erg g^{-1} yr^{-1}.
[b]Temporary increase in electrical conductivity during irradiation at dose rates $\sim 10^{8}$ erg g^{-1} yr^{-1}.

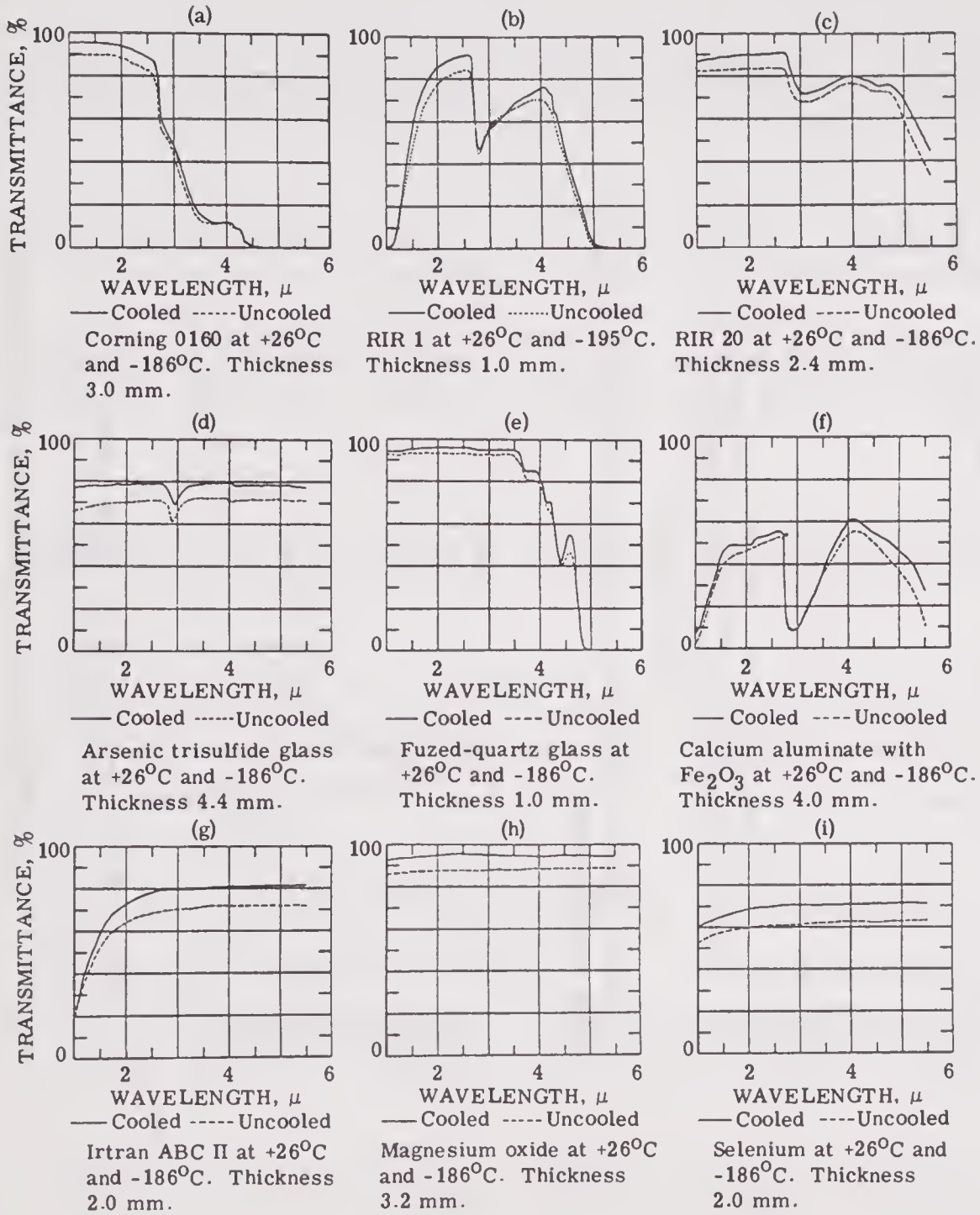


FIG. 8-38. Transmittance of cooled optical materials.

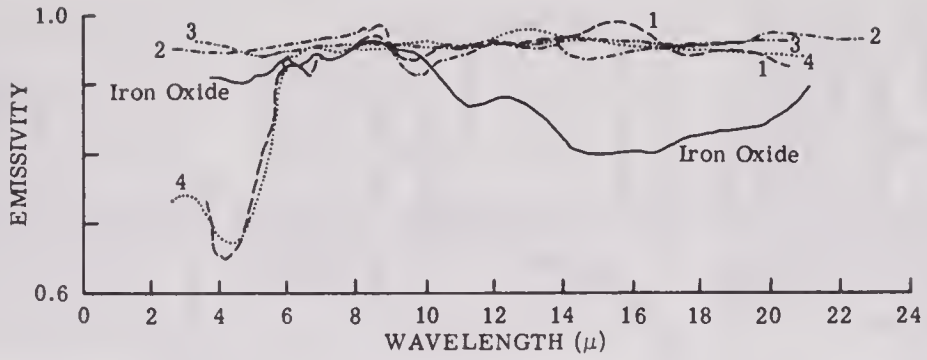


FIG. 8-39. Specular emittance of some useful "blacks;" the samples are from Barnes Engineering Co. and identified both by their Barnes label and their generic name: 1. Pryomark Standard Black (BEC-1); 2. Sicon Black (BEC-2); 3. Krylon Flat Black (BEC-3); 4. Magnesium oxide powder (BEC-4); 5. Iron oxide.

8.10. Optical Properties of Water

Water in many forms – vapor, liquid, solid, distilled, brine, etc. – has a number of applications in the infrared. Some of the useful and interesting properties are given here. Normal incidence values of specular reflectivity and values of n and k are given by Centeno [75] are given in Table 8-21; Fig. 8-40 gives the reflection and Fig. 8-41 the refractive index. Values for different angles of incidence are given by McSwain and Bernstein [76]. Similar data are given by Kislovskii [77]; Ockman [78] gives information on ice.

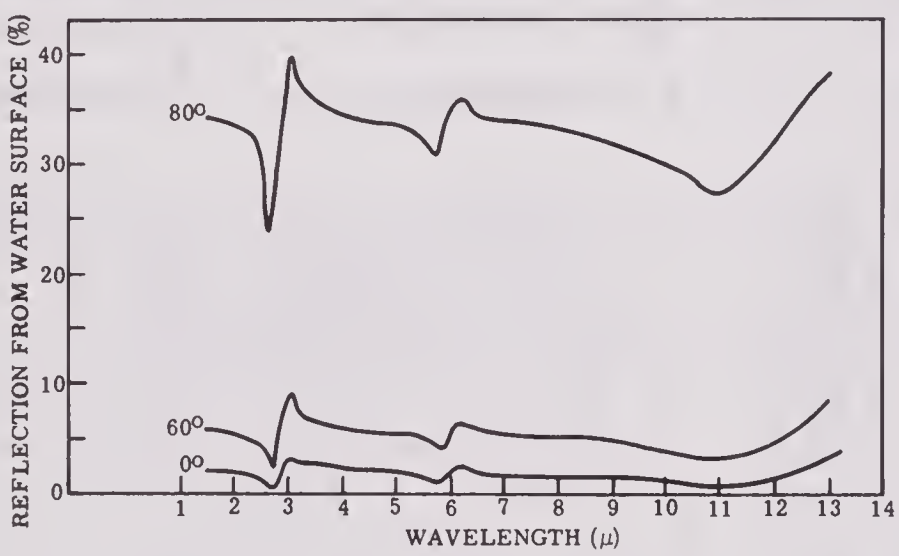


FIG. 8-40. Specular reflectance of water as a function of incident angle and wavelength.

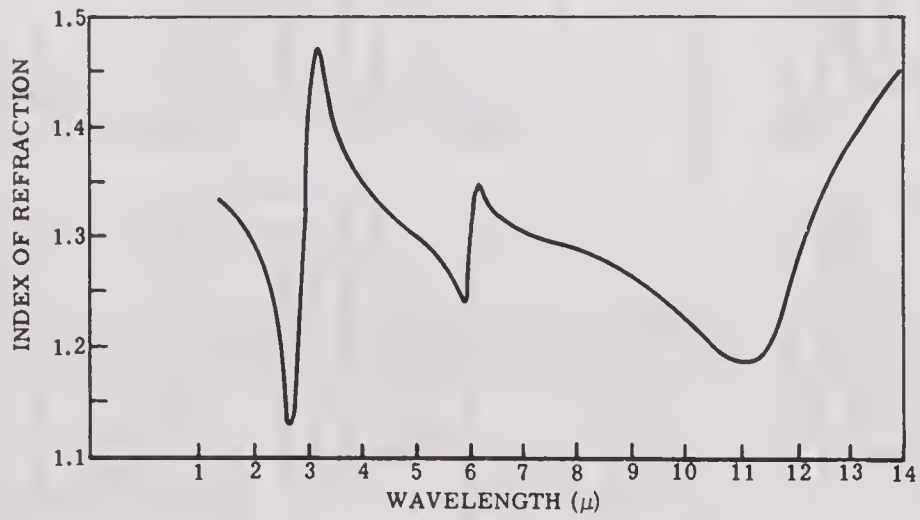


FIG. 8-41. Index of refraction of water calculated from reflectivity.

TABLE 8-21. OPTICAL CONSTANTS OF WATER IN THE NEAR INFRARED [75]

Wavelength (μ)	*Reflectivity* (%)	*Extinction Coefficient*	*Refractive Index*	*Wavelength* (μ)	*Reflectivity* (%)	*Extinction Coefficient*	*Refractive Index*
1.00	1.96	3.5×10^{-6}	1.326	5.85	1.25	0.0653	1.242
1.05	1.95	3.1	1.325	5.90	1.50	0.0799	1.266
1.10	1.94	1.7	1.324	6.00	2.02	0.1220	1.304
1.20	1.93	1.21×10^{-5}	1.323	6.04	2.16	0.1372	1.312
1.30	1.91	1.29	1.321	6.10	2.28	0.1216	1.331
1.40	1.90	1.30×10^{-4}	1.320	6.20	2.46	0.0940	1.358
1.50	1.88	3.12	1.318	6.30	2.34	0.0768	1.352
1.60	1.87	1.12	1.316	6.40	2.22	0.0626	1.344
1.70	1.85	1.08	1.315	6.50	2.12	0.0524	1.336
1.80	1.82	1.43	1.312	6.60	2.05	0.0462	1.331
1.90	1.79	8.14	1.309	6.70	2.00	0.0450	1.326
2.00	1.74	14.3	1.304	6.80	1.98	0.0448	1.324
2.20	1.63	5.11	1.293	6.90	1.97	0.0451	1.323
2.40	1.47	10.1	1.276	7.00	1.95	0.0457	1.321
2.50	1.37	20.1	1.265	7.20	1.87	0.0463	1.313
2.60	1.25	51.8	1.252	7.40	1.77	0.0466	1.303
2.70	0.96	0.0183	1.216	7.60	1.72	0.0484	1.298
2.74	0.75	0.0273	1.187	7.80	1.68	0.0472	1.294
2.77	0.90	0.0364	1.206	8.00	1.66	0.0472	1.292
2.80	1.41	0.0490	1.266	8.20	1.64	0.0472	1.289
2.85	2.00	0.0578	1.324	8.40	1.62	0.0473	1.287
2.90	2.48	0.0637	1.367	8.60	1.57	0.0476	1.282
2.95	2.87	0.0670	1.401	8.80	1.51	0.0483	1.276
3.00	3.40	0.0680	1.446	9.00	1.44	0.0498	1.268
3.02	3.90	0.0683	1.453	9.2	1.35	0.0518	1.257
3.07	4.39	0.0682	1.525	9.4	1.27	0.0542	1.247
3.10	4.30	0.0681	1.517	9.6	1.18	0.0562	1.236
3.16	4.12	0.0658	1.504	9.8	1.09	0.0595	1.224
3.20	4.00	0.0611	1.495	10.0	0.99	0.0601	1.212
3.30	3.65	0.0370	1.471	10.2	0.92	0.0621	1.202
3.40	3.37	0.0225	1.449	10.4	0.85	0.0687	1.190
3.50	3.05	0.0128	1.423	10.6	0.78	0.0802	1.175
3.60	2.80	0.00783	1.402	10.8	0.73	0.0946	1.159
3.70	2.59	0.00619	1.381	10.9	0.70	0.0993	1.150
3.80	2.36	0 00596	1.363	11.0	0.72	0.1138	1.143
3.90	2.25	0.00590	1.353	11.1	0.77	0.1290	1.137
4.03	2.19	0.00642	1.347	11.2	0.84	0.1471	1.129
4.10	2.18	0.00709	1.346	11.3	0.92	0.1637	1.121
4.20	2.16	0.00871	1.344	11.4	0.97	0.1751	1.114
4.30	2.15	0.0108	1.343	11.5	1.02	0.1831	1.111
4.40	2.14	0.0133	1.342	11.6	1.09	0.1880	1.118
4.50	2.13	0.0164	1.341	11.7	1.19	0.1937	1.130
4.60	2.11	0.0194	1.339	11.8	1.30	0.1991	1.144
4.70	2.10	0.0202	1.338	12.0	1.47	0.2058	1.165
4.80	2.08	0.0199	1.336	12.5	2.15	0.2438	1.219
4.90	2.05	0.0188	1.334	13.0	3.02	0.2918	1.270
5.00	2.02	0.0169	1.331	13.5	3.55	0.3202	1.297
5.10	1.98	0.0152	1.327	14.0	4.10	0.3582	1.309
5.20	1.93	0.0143	1.322	14.5	4.72	0.4028	1.313
5.30	1.86	0.0143	1.315	15.0	5.12	0.4298	1.315
5.40	1.80	0.0146	1.309	16.0	5.29	0.4088	1.374
5.50	1.73	0.0174	1.302	17.0	6.15	0.4070	1.468
5.60	1.66	0.0232	1.295	18.0	6.21	0.4585	1.401
5.70	1.57	0.0340	1.284				
5.80	1.40	0.0515	1.263				

8.11. Effects of Space Radiation on Optical Materials

The large flux and energy of the electrons in the low-lying Van Allen belt indicates that effects on the spectral transmission of various materials may be appreciable. An example of the effect of 1-Mev electrons on the transmission of various materials is shown in Table 8-22, taken from data by Cooley and Janda [79]. Many of the materials shown in Table 8-22 are also used as cover materials for solar cells in satellites.

The decrease in transmission of fused quartz windows as a result of radiation occurs fairly rapidly during the initial exposure, with the change being much less in magnitude as the total radiation dose accumulates. There have also been large differences observed in the amount of reduction in transmission as a result of radiation, and this is believed to be due to the impurity content of the quartz. In the studies of solar cell windows [79] the poorest quartz had 50 to 70 ppm of impurities, and the best had less than 1 ppm. From these studies it was apparent that extreme care is essential in the selection of transparent material such as fused quartz if a minimum of reduced transmission as a result of radiation is desired.

In the case of glasses and crystals used as optical elements in infrared systems, the effects of radiation have not been extensively studied. In the case of glasses it would be expected that color formation under a radiation environment would be severe. As a result of "built in" impurity centers in the glass network, radiation may form displaced electrons with negative ion sites or vacancies, to form color centers. It appears, however, that although color centers severely affect the visible transmission, in some cases they do not appreciably influence the transmission characteristics in the infrared region. Until more information is available it appears that the problems for a given optical element might best be solved by exposing this element to radiation levels corresponding with the flux and energy expected to be present in the particular orbit encountered in its space environment. The element should be examined before and after radiation in the desired spectral region.

TABLE 8-22. EFFECTS OF 1-MEV ELECTRONS ON SPECTRAL TRANSMITTANCE OF OPTICAL MATERIALS (TOTAL DOSE OF 10^{16} e/cm^2)

Material Description		*50% Cutoff Point (μ)*	*Wavelength (μ)* 0.40	0.50	0.60	0.70
1. Microsheet, Corning 0211, 6 mil	T_0	–	89.0	90.0	90.0	90.0
	T_F	–	82.0	85.0	87.0	99.0
	ΔT	–	7.9	5.6	3.3	2.2
2. Fused silica, Corning 7940, 66 mil	T_0	–	89.0	90.0	90.0	90.0
	T_F	–	88.5	88.5	88.0	89.0
	ΔT	–	0.6	1.7	2.2	1.1
3. Fused silica, Corning 7940, 20 mil + antireflecting coating + "blue" filter	T_0	0.416	Below 50% cutoff	89.0	90.0	91.0
	T_F	0.425	Below 50% cutoff	88.0	88.0	90.0
	ΔT	–	–	1.1	2.2	1.1
4. Nonbrowning lime glass Corning 8365, density 2.7, 60 mil*	T_0	0.365	89.0	89.0	89.0	89.0
	T_F	–	29.0	30	33	35
	ΔT	–	–	–	–	–
5. Nonbrowning lead glass Corning 8365, density 3.3, 60 mil	T_0	0.380	72.0	73.0	74.0	74.0
	T_F	–	21.0	22.0	22.0	24.0
	ΔT	–	–	–	–	–
6. High-density lead glass	T_0	0.392	82.0	84.0	85.0	84.0
	T_F	0.402	72.0	78.0	82.0	83.0
	ΔT	–	12.2	7.1	3.5	1.2

T_0 = Initial preirradiated transmittance.
T_F = Transmittance after 10^{16} e/cm^2.
ΔT = Decrease in transmittance.
*Radiation caused internal crazing in specimen diffusing light.

References

1. S. S. Ballard, K. A. McCarthy, and W. L. Wolfe, *Optical Materials for Infrared Instrumentation*, 2389-11-S (1959); Supplement 2389-11-S_1 (1961), The University of Michigan, Ann Arbor.
2. J. Jerger, Jr., *Investigation of Long Wavelength Infrared Transmitting Glasses,* Tech. Document Report ASD-TDR-63-552, Servo Corp. of America, Hicksville, N.Y. (1963).
3. *New High Temperature Infrared Transmitting Glasses,* Texas Instruments Incorporated, Dallas, Texas (1963).
4. M. Herzberger and C. D. Salzberg, *J. Opt. Soc. Am.*, **52**, 420 (1962).
5. *Quarterly Progress Report, July, August and September 1963,* Tech. Rept. No. 2900-455-P, Institute of Science and Technology, The University of Michigan, Ann Arbor, Mich. (Nov 1963). CONFIDENTIAL
6. A. E. Conrady, *Applied Optics and Optical Design*, Dover, New York (1957).
7. O. S. Heavens, *Optical Properties of Thin Solid Films*, Butterworth Scientific Publications, London (1955).
8. H. O. McMahon, *J. Opt. Soc. Am.*, **40**, 376 (1950).
9. G. Hass and A. F. Turner, "Coatings for Infrared Optics," Reprint from Wissenschaftliche Verlagsgesellschaft m. b. h., Stuttgart, 143-163.
10. Private Communication, Haynes Stellite Co., Kokomo, Ind.
11. J. R. Jenness, Jr., *J. Opt. Soc. Am.*, **46**, 157 (1956).
12. J. T. Cox and G. Hass, *J. Opt. Soc. Am.*, **48**, 677 (1958).
13. A. P. Bradford, W. W. Erbe, and G. Hass, *J. Opt. Soc. Am.*, **49**, 990-991 (1959).
14. G. R. Hennig, Moderator, "Shields and Auxiliary Equipment," talk given at the colloquium on the Effect of Radiation on Materials at Johns Hopkins University, Baltimore, Md. (1957).
15. T. A. Longo, *Nucleon Irradiation of Silicon Semiconductors*, Special Report to Signal Corps., Purdue University, LaFayette, Ind. (1957).
16. L. D. Jaffe and J. B. Rittenhouse, *Behavior of Materials in Space Environments*, Technical Report 32-150, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, Calif. (1961).
17. R. W. King, N. J. Broadway, and S. Palinchak, *The Effect of Nuclear Radiation on Elastomeric and Plastic Components and Materials*, REIC Report No. 21, The Radiation Effects Information Center, Battelle Memorial Inst., Columbus, Ohio (1961).
18. K. Kobayaski, *Phys. Rev.*, **102**, 348 (1956).
19. A. F. Cohen and L. C. Templeton, *Solid State Division Progress Report*, ORNL-2188, Oak Ridge National Laboratories, Oak Ridge, Tenn. (1956).
20. R. Berman, F. E. Simon, P. G. Klemens, and T. M. Fry, *Nature*, **166**, 277 (1950).
21. J. J. Harwood, H. H. Hausner, and J. G. Morse, *The Effects of Radiation on Materials,* Reinhold, New York (1958).
22. F. A. Bovey, *Effects of Ionizing Radiation on Natural and Synthetic High Polymers*, Interscience, New York (1958).
23. N. J. Broadway, M. A. Youtz, S. Palinchak, and R. Mayer, *Effect of Nuclear Radiation on Elastometer and Plastic Materials*, Battelle Memorial Institute, Radiation Effects Information Center, Rept. 3 and Addenda, Columbus, Ohio (1958-1960).
24. J. F. Fowler and F. T. Farmer, *Nature*, **171**, 1020-1021 (1953).
25. J. F. Fowler and F. T. Farmer, *Nature*, **174**, 136-137 (1954).
26. L. A. Wall and R. E. Florin, *J. Appl. Polymer. Sci.* **1**, 251 (1959).
27. C. D. Bopp and O. Sisman, *Nucleonics*, **13**, No. 7, 28-33 (1955).
28. C. D. Bopp and O. Sisman, *Nucleonics*, **13**, No. 10, 51-55 (1955).
29. R. A. Mayer, N. J. Broadway, and S. Palinchuk, *Effect of Nuclear Radiation on Protective Coatings*, Battelle Memorial Institute, Radiation Effects Information Center, Rept. 13, Columbus, Ohio (1960).
30. J. P. Shoffner, *J. of Teflon*, **2**, No. 1, 6-7 (1961).
31. J. H. Coleman and D. Bohm, *J. Appl. Phys.*, **24**, 497-498 (1953).
32. A. J. Warner, F. A. Muller, and H. G. Nordlin, *J. Appl. Phys.*, **25**, 131 (1954).
33. P. Y. Feng and J. W. Kennedy, *J. Am. Chem. Soc.*, **77**, Part 1, 847-851 (1955).
34. A. Charlesby, "Effects of Radiation on Behavior and Properties of Polymers," in J. J. Harwood, H. Hausner, J. G. Morse, and W. G. Rauch (eds.), *Effects of Radiation on Materials*, Reinhold, New York, 261-286 (1958).
35. K. R. Ferguson, "Design and Construction of Shielding Windows," *Nucleonics 10*, No. 11, 46-51 (1952).
36. G. S. Monk, "Coloration of Optical Glass by High-Energy Radiation," *Nucleonics 10*, No. 11, 52-55 (1952).

37. N. J. Kreidl and J. R. Hensler, *J. Am. Ceram. Soc.*, **38**, 423-32 (1955).
38. N. J. Kreidl and J. R. Hensler, *J. Opt. Soc. Am.*, **47**, 73-74 (1957).
39. R. L. Hines, *J. Appl. Phys.*, **28**, 587-91 (1957).
40. J. F. Hansen, S. E. Harrison, W. L. Hood, D. J. Hamman, W. E. Chapin, and E. N. Wyler, *Effect of Nuclear Radiation on Electronic Components and Systems,* Battelle Memorial Institute, Radiation Effects Information Center, Repts. 2 and 12, Columbus, Ohio (1957 and 1960).
41. J. K. Davis, *Electrical Mfg. 59*, No. 6, 151-156 (1957).
42. M. Levy and J. H. O. Varley, *Proc. Phys. Soc.*, **B68**, 223-233 (1955).
43. A. J. Gale and F. A. Bickford, *Nucleonics 11*, No. 8, 48 (1953).
44. F. S. Dainton and J. Rowbottom, *Trans. Faraday Soc.*, **50**, 480-493 (1954).
45. P. W. Levy, *J. Chem. Phys.*, **23**, 764-765 (1955).
46. W. Primak, L. H. Fuchs, and P. Day, *J. Am. Ceram. Soc.*, **38**, 135-139 (1955).
47. J. H. Crawford, Jr., and M. C. Wittels, "Review of Investigations of Radiation Effects in Ionic and Covalent Crystals," *Proc. Intl. Conf. on Peaceful Uses of Atomic Energy*, United Nations, N.Y. (1955), **7**, Nuclear Chemistry and Effects of Irradiation, 654-665.
48. M. Swerdlow and R. F. Geller, *Survey of Radiation-Resistant Glass*, U.S. National Bureau of Standards unpublished report to U.S. Air Force, Wright Air Development Center, on Order 33(616)56-17 (1957).
49. L. Reiffel, R. Estin, D. Kazen, L. Marchi, and H. Nakamura, *Radiation Sensitive Glasses*, Armour Research Foundation, Final Rept. to U.S. Army Evans Signal Lab., Project A031-4, Chicago (1953).
50. C. M. Nelson and R. A. Weeks, *J. Appl. Phys.*, **32**, 883-886 (1961).
51. E. W. J. Mitchell and E. G. S. Paige, *Phil. Mag.*, Series 8, **1**, 1085-1115 (1956).
52. G. J. Dienes and G. H. Vineyard, *Radiation Effects in Solids,* Interscience, New York (1957).
53. F. Seitz and J. S. Koehler, *Displacement of Atoms During Irradiation*, in F. Seitz and D. Turnbull (eds.), Academic Press, New York, **2**, 307-448 (1956).
54. D. S. Billington, "Radiation Effects in Metals and Alloys," in *Effects of Radiation on Materials*, J. J. Harwood, H. Hausner, J. G. Morse, and W. G. Rauch (eds.), Reinhold, New York, 99-125 (1958).
55. B. C. Allen, A. K. Wolff, A. R. Elsea, and P. D. Frost, *Effect of Nuclear Radiation on Structural Metals*, Battelle Memorial Inst. Radiation Effects Information Center, Rept. 5, Columbus, Ohio (1958).
56. P. J. Reid, J. W. Moody, *Effect of Nuclear Radiation on Magnetic Materials,* Battelle Memorial Inst. Radiation Effects Information Center, Tech. Memo 12, Columbus, Ohio (1958).
57. J. M. Denney, "Radiation Damage to Satellite Solar Cell Power Systems." *Am. Rocket Soc. Reprint* 1295-60 (1960); in "Energy Conversion for Space Power," ed. by N. W. Snyder, Academic Press, New York, 345-61 (1961).
58. H. Y. Fan and K. Lark-horovitz, "Irradiation Effects in Semiconductors," in *Effects of Radiation on Materials,* J. J. Harwood, H. Hausner, J. G. Morse, and W. G. Rauch (eds.), Reinhold, New York, 99-125 (1958).
59. C. M. Nelson and R. A. Weeks, *J. Appl. Phys.*, **32**, 883-886 (1961).
60. J. C. Pigg and C. G. Robinson, *Electrical Mfg.*, **59**, No. 4, 116-124 (1957).
61. G. Enslow, F. Junga, and W. W. Happ, *Gamma Radiation Effects in Silicon Solar Cells*, in 3rd ARDC Radiation Effects Symp., 4, Lockheed Nuclear Products Div., Marietta, Georgia (1958).
62. G. L. Keister and H. V. Steward, *Proc. Inst. Radio Eng.*, **45**, 931-937 (1957).
63. H. L. Steele, "Effects of Gamma Radiation on Transistor Parameters," *Proc. Transistor Reliability Symp.*, New York University, New York, 96 (1956).
64. J. E. Drennen and S. E. Harrison, *Effect of Nuclear Radiation on Semiconductor Diodes*, Battelle Memorial Institute, Radiation Effects Information Center Tech. Memo 6, Columbus, Ohio (1958).
65. F. Gordon, *Effects of Nuclear Radiation on Power Transistors*, in 3rd ARDC Radiation Effects Symposium, 4, Lockheed Aircraft Corp., Nuclear Products Div., Marietta, Georgia (1958).
66. T. E. Lusk, *Electronic Design*, 8, No. 22, 74-75 (1960).
67. R. G. Downing, "Electron Bombardment of Silicon Solar Cells," *A. Rocket Soc. Reprint 1294-60* (1960); *Energy Conversion for Space Power*, N. W. Snyder (ed.), Academic Press, New York, 325-344 (1961).
68. D. S. Billington, "Radiation Effects in Metals and Alloys," *Effects of Radiation on Materials,* J. J. Hardwood, H. Hausner, J. G. Morse, and W. G. Rauch (eds.), Reinhold, New York, 99-125 (1958).

69. B. C. Allen, A. K. Wolff, A. R. Elsea, and P. D. Frost, *Effect of Nuclear Radiation on Structural Metals,* Battelle Memorial Inst. Radiation Effects Information Center, Rept. 5, Columbus, Ohio (1958).
70. F. J. Reid and J. W. Moody, *Effect of Nuclear Radiation on Magnetic Materials*, Battelle Memorial Inst., Radiation Effects Information Center, Tech. Memo 12, Columbus, Ohio (1958).
71. G. F. Linsteadt and H. P. Leet, *Proc. IRIS*, **6**, No. 3, 159 (1961).
72. D. L. Stierwalt, D. D. Kirk, and J. B. Bernstein, *Foundational Research Projects,* July-September 1962, Navweps Report 7237, Naval Ordnance Laboratory, Corona, Calif. (1962). Also in in *Appl. Optics,* **2**, 1169 (1963).
73. L. Harris and K. F. Cuff, *J. Opt. Soc. Am.*, **46**, 160 (1956).
74. Henry H. Blau, Jr., Eleanor Chaffee, John R. Jasperse, and William S. Martin, *High Temperature Thermal Radiation Properties of Solid Materials,* Arthur D. Little, Inc., Cambridge, Mass. (1960).
75. M. Centeno, *J. Opt. Soc. Am.*, **31**, 245 (1941).
76. B. McSwain and J. Bernstein, *Foundational Research Projects*, October-December 1960, NOLC 539, Naval Ordnance Laboratory, Corona, Calif. (1961).
77. L. D. Kislovskii, *Optics and Spectroscopy*, **7**, 201 (1959).
78. N. Ockman, Dissertation, University of Michigan, Ann Arbor (1957).
79. W. C. Cooley and R. J. Janda, *Handbook of Space Radiation Effects on Solar Cell Power Systems,* NASA Rept. SP-3003 (1963).

Chapter 9

OPTICS

Warren J. Smith

Infrared Industries, Inc.

CONTENTS

9.1. Terminology, Symbols, and Sign Convention 371
9.1.1. Terminology 371
9.1.2. Symbols 371
9.1.3. Sign Convention 371
9.2. First-Order (Gaussian) Optics 371
9.2.1. Focal Points and Principal Points 371
9.2.2. The Cardinal Points of Elements 373
9.2.3. Multielement Optical Systems 375
9.2.4. Paraxial Ray Trace – Surface by Surface 377
9.3. Limitation of Rays by Stops and Apertures 377
9.3.1. The Aperture Stop 377
9.3.2. The Field Stop 378
9.3.3. Vignetting 378
9.3.4. Glare Stops and Baffling 379
9.3.5. Relative Aperture, Speed, F/No, Numerical Aperture 379
9.3.6. Depth of Field and Depth of Focus 380
9.4. Aberrations 381
9.4.1. Aberration Description 381
9.4.2. The Seidel Aberrations 382
9.4.3. Correction of Aberrations 385
9.4.4. Variations of Aberrations with Aperture and Image Size 385
9.4.5. Zonal and Residual Aberrations 386
9.4.6. Chromatic Variation of Aberrations 386
9.4.7. Graphical Representation of Aberrations 386
9.5. Ray Tracing 388
9.5.1. Ray-Tracing Precision 388
9.5.2. Determination of Specific Aberrations 388
9.5.3. Ray-Tracing Equations 390
9.5.4. Graphical Ray Tracing 398

9.6. Third-Order Aberrations 399
9.6.1. Third- and Higher-Order Aberrations – Surface Contributions 399
9.6.2. Third-Order Aberration Contributions of Thin Lenses 402
9.6.3. Stop Shift Theory 404
9.6.4. Afocal Systems 406
9.7. Optical Design Techniques 406
9.7.1. General Considerations 406
9.7.2. Correction of Primary Aberrations 406
9.7.3. Reduction of Residual Aberrations 407
9.7.4. Automatic Design. 408
9.8. Achromatism and Achromats 409
9.9. Resolution, Definition, and Image Spot Size 410
9.9.1. The Airy Disc 410
9.9.2. The Effects of Aberrations on the Airy Disc 413
9.9.3. Geometrical Limits on Resolution 413
9.10. Summary of Equations 417

9. Optics

9.1. Terminology, Symbols, and Sign Conventions

The formulas of geometrical optics are the result of selective application of Snell's law,

$$n' \sin I' = n \sin I \tag{9-1}$$

This can also be written in vector form. Let $\mathbf{Q}_0$ be a vector representing the incident ray and $\mathbf{Q}_1$ the emergent ray, and let $\mathbf{N}$ = the surface normal. Then Snell's law is

$$(n_0\mathbf{Q}_0 - n_1\mathbf{Q}_1) \times \mathbf{N} = 0 \tag{9-2}$$

Optical systems are usually composed of a series of surfaces of revolution, often spherical. The optical axis of a lens or system is that line upon which the centers of curvature of the surfaces lie. Unless the contrary is specifically noted, optical systems are assumed to be axially symmetrical.

9.1.1. Terminology.

An *element* is a single piece of optical material (*e.g.*, a simple lens or mirror).

A *component* is one or more elements treated as a unit (*e.g.*, a cemented doublet).

A *member* consists of all the components either ahead of (the front member) or behind (the rear member) the aperture stop.

9.1.2. Symbols. In general, lower-case symbols are used for the paraxial (see Sec. 9.2) values of quantities and capitals, for values other than paraxial. Primed symbols refer to quantities after refraction by a surface or a lens. Subscripts identify the surface or lens with which a quantity is associated.

9.1.3. Sign Convention. With the exception of certain aberrations, the sign convention follows that given in *Applied Optics* and *Optical Design*, by A. E. Conrady, Dover, New York (1957). Light is assumed to progress from left to right. Radii and curvatures are positive if the center of curvature is to the right of the surface. Surfaces or elements are positive if they converge light. Points lying to the right of a point, element, or surface are considered to be a positive distance away. Slope angles are positive if the ray must be rotated counterclockwise to reach the axis. Angles of incidence and refraction are positive if the ray must be rotated clockwise to reach the normal to the surface.

9.2. First-Order (Gaussian) Optics

The first-order expressions are derived by reducing the exact trigonometrical expressions for ray paths to the limit when the angles and ray heights involved approach zero. These expressions, however, are completely accurate for the paraxial region.

9.2.1. Focal Points and Principal Points. A well-corrected optical system can be treated as a "black box" whose characteristics are defined by its cardinal points, which are its first and second focal points and its first and second principal points. The cardinal points are shown in Fig. 9-1.

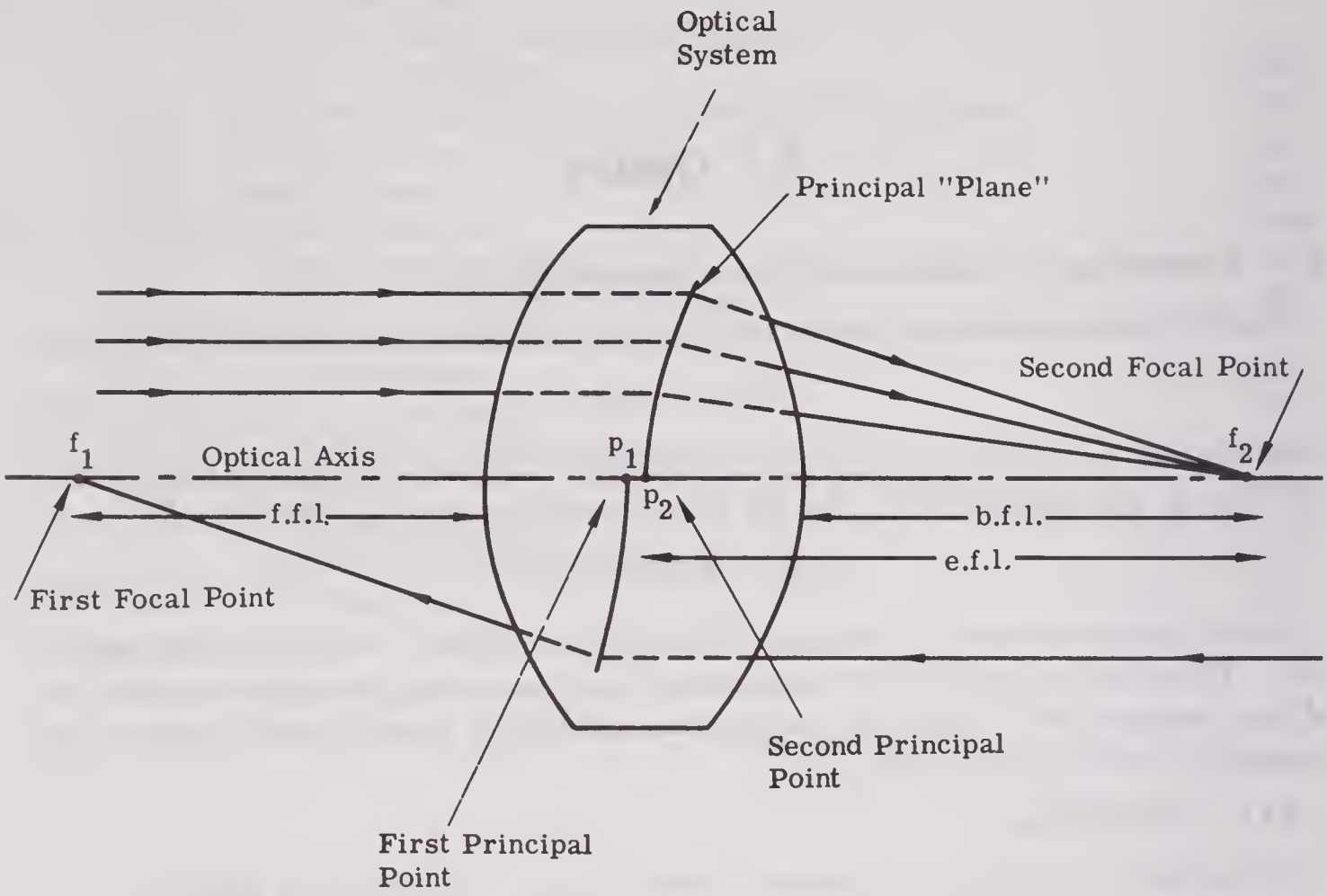


FIG. 9-1. Location of focal and principal points of an optical system.

9.2.1.1. *Image Position.* The location of the image formed by an optical system can be determined by either of the following expressions:

$$\frac{1}{s'} = \frac{1}{s} + \frac{1}{f}; \qquad f^2 = -xx' \tag{9-3}$$

(the quantities are defined in Fig. 9-2). If the object or image lies to the left of the principal or focal point, s, s', x, and x' are negative.

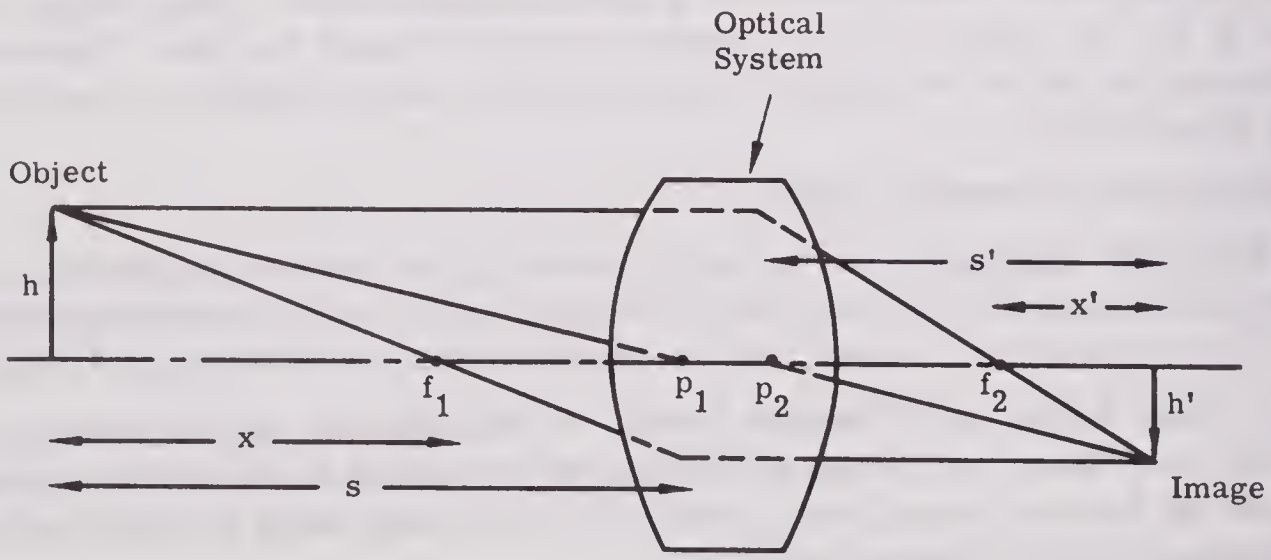


FIG. 9-2. Relationships between cardinal points and position and size of object and image.

9.2.1.2. *Magnification.* Lateral magnification relates the heights of image and object and is given by

$$m = -f/x = -x'/f = h'/h = s'/s \tag{9-4}$$

(the quantities are defined in Fig. 9-2).

Longitudinal magnification relates the lengths (or depths) along the optical axis of image and object and is given by

$$\overline{m} = s'_1 s'_2 / s_1 s_2 = m^2 \text{ (approx)} \tag{9-5}$$

where s_1 is the distance from the principal points to the edges of the object and image nearest the optical system, and s_2 is the distance from the principal points to the edges most distant from the optical system.

9.2.1.3. *The Lagrangian Invariant.* Figure 9-3 shows an axial ray and a principal ray passing through an image-forming optical system. The Lagrangian invariant applies wherever either ray crosses the axis.

$$I = hnu = h'n'u' = ynu_{pr} \tag{9-6}$$

Outside the paraxial region the invariant takes the form hn (sin u), and is also called the Law of Sines. A rearrangement of Eq. (9-6) yields an additional relationship for the magnification:

$$m = \frac{h'}{h} = \frac{nu}{n'u'} = \frac{u}{u'} \text{ (in air)} \tag{9-7}$$

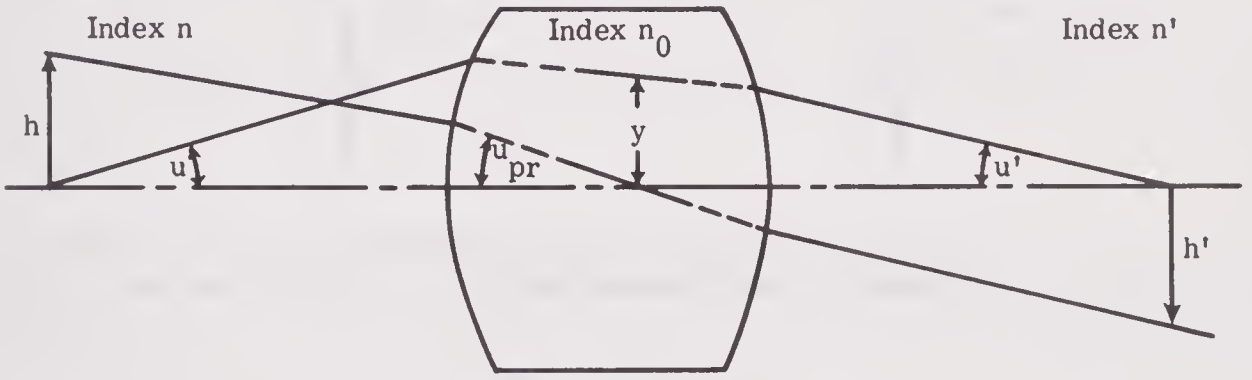


FIG. 9-3. The Lagrangian invariant.

9.2.2. The Cardinal Points of Elements. Figure 9-4 shows the location of the cardinal points of different types of elements. Their values are (see Fig. 9-4*a* for definitions of symbols):

$$f = -f' = -\frac{nr_1r_2}{\Delta} \qquad \begin{matrix} \textit{convergent } f > 0 \\ \textit{divergent} \quad f < 0 \end{matrix}$$

$$\delta = n\frac{r_1^2}{\Delta}, \qquad \delta' = -n\frac{r_2^2}{\Delta}$$

$$d = (n-1)\frac{r_1 t}{\Delta}, \qquad d' = (n-1)\frac{r_2 t}{\Delta}$$

where $\Delta = (n-1)[n(r_1 - r_2) - (n-1)t]$

$$f_0 = \frac{r_1}{n-1}, \qquad f'_0 = -\frac{nr_1}{n-1}$$

$$f_1 = \frac{nr_2}{1-n}, \qquad f'_1 = -\frac{r_2}{1-n}$$

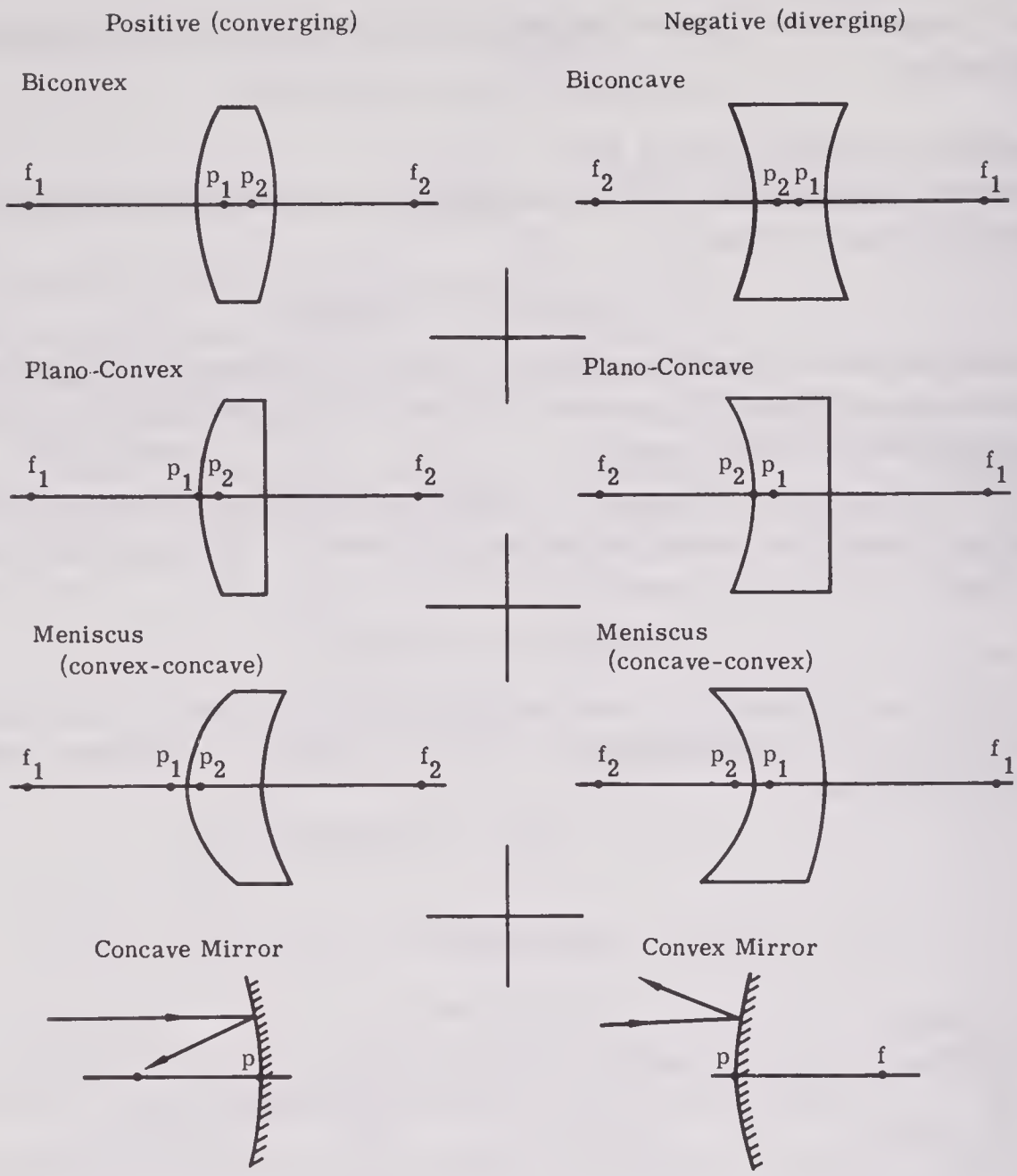


FIG. 9-4. Location of principal points for thick lenses of various common shapes.

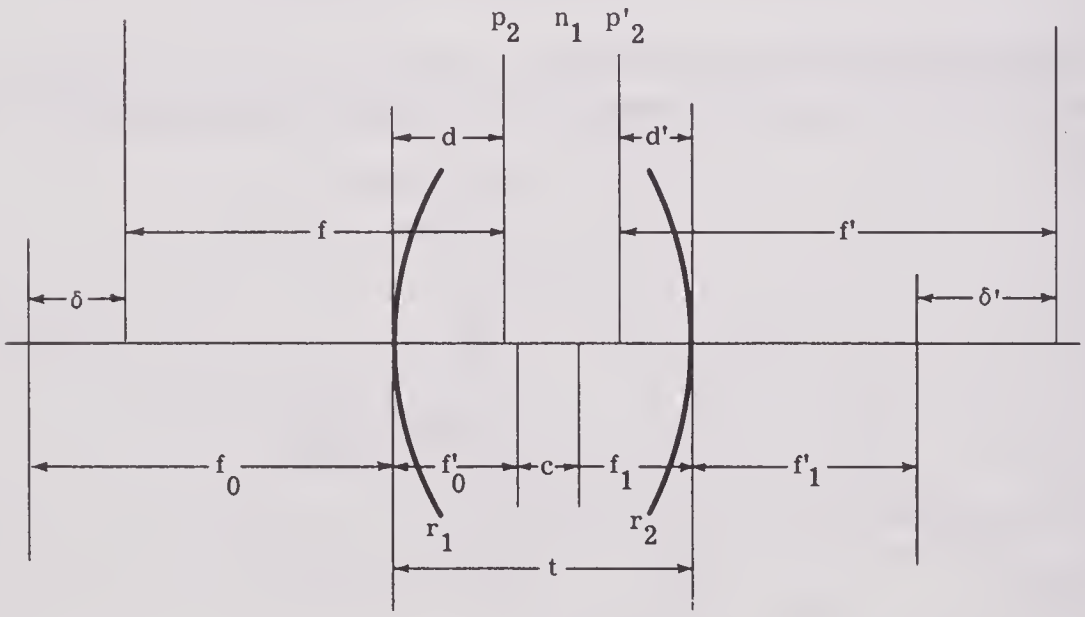


FIG. 9-4*a*. Definitions of symbols.

9.2.2.1. *Thick-Lens Equations.* The cardinal points of an element of appreciable thickness are given by the following equations:

$$\text{power} = \phi = \frac{1}{f} = (n-1)\left[\frac{1}{r_1} - \frac{1}{r_2} + \frac{(n-1)t}{nr_1r_2}\right]$$

$$= (n-1)\left[c_1 - c_2 + \frac{(n-1)}{n} tc_1c_2\right] \tag{9-8}$$

$$\text{back focus} = \text{b.f.l.} = f - \frac{(n-1)}{nr_1} tf \tag{9-9}$$

9.2.2.2. *Thin-Lens Equations.* A thin lens can be defined as a lens whose thickness is negligible compared to the accuracy required for the calculation. If one sets $t = 0$, Eq. (9-8) reduces to

$$\phi = \frac{1}{f} = (n-1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) = (n-1)(c_1 - c_2) \tag{9-10}$$

The power of a thin element can be held constant for an infinite number of element shapes by varying the ratio of the curvatures of radii while $(c_1 - c_2)$ is held constant, a process called bending.

The principal points of a thin lens are taken to be coincident with its location.

9.2.2.3. *Mirrors.* The cardinal points of a mirror in air may be derived by assuming an index of (-1) for the medium following the reflecting surface and applying the thin-lens expression for power, neglecting c_2. Thus, for a mirror,

$$\phi = \frac{1}{f} = (-1-1)\frac{1}{r} = \frac{-2}{r} = -2c \tag{9-11}$$

$$f = -0.5r$$

The principal points are coincident with the mirror surface.

The sign convention for distance after reflection is reversed as a result of the reflection index of (-1). Thus, the distance from the mirror to an image lying to the left of the mirror is taken as positive instead of negative.

9.2.3. Multielement Optical Systems. For optical systems of more than two elements, it is preferable to determine the image locations and sizes, and the cardinal points, by tracing the paths of specific rays. It is also possible to calculate the position and size of an image by repeated application of Eqs. (9-3), (9-4), and (9-5).

9.2.3.1. *Ray Tracing a Series of Elements.* Given a series of elements with powers $\phi_1, \phi_2, \phi_3, \ldots, \phi_n$ separated by distances $d'_1, d'_2, d'_3, \ldots, d'_{n-1}$, the path of a ray through the series may be traced by repeated applications of the following equations:

$$u'_i = u_i + y_i\phi_i \tag{9-12}$$

$$y_{i+1} = y_i - d'_i u_i \tag{9-13}$$

where u_i is the slope of the ray approaching the element i

u'_i is the slope of the ray it has passed after through element i

y_i is the height at which the ray strikes the element

(These equations are most useful for tracing through a series of thin lenses, but they can be applied equally well to thick lenses by taking y as the height at which the ray strikes the principal planes and d as the spacing between principal points.)

Thus, to determine the cardinal points of a system, one can trace a ray with $u_1 = 0$ and a nominal value for y_1. Then the effective focal length is given by

$$f = y_1/u'_n \tag{9-14}$$

and the back focal length is given by

$$b = y_n/u'_n \tag{9-15}$$

Reversing the system and repeating the process will yield the other set of cardinal points.

Alternatively, a ray may be traced from the axial intercept of the object through the system. The starting slope for such a ray is given by

$$u_1 = y_1/s_1 \tag{9-16}$$

and the image is located at

$$s'_n = y_n/u'_n \tag{9-17}$$

The size of the image is then

$$h' = hu_1/u'_n \tag{9-18}$$

Also, a principal ray may be traced from a specific off-axis point in the object; its intercept in the image plane is then the image of the object point. This height may thus be found from

$$h' = y_{pr} - s'u'_{pr} \tag{9-19}$$

where the subscript pr indicates the principal-ray data.

9.2.3.2. *Combination of Two Elements.* Given two elements, A and B, separated by a distance d, with focal lengths f_A and f_B, powers ϕ_A and ϕ_B, a combined focal length f_{AB}, and back focus s', and v the distance from element A to the focus ($v = d + s'$), then one or more of the following expressions will be found to cover almost any desired relationship.

$$f_{AB} = \frac{f_A f_B}{f_A + f_B - d} \tag{9-20}$$

$$s' = \frac{f_B(f_A - d)}{f_A + f_B - d} = f_{AB}\frac{(f_A - d)}{f_A} = f_{AB}(1 - \phi_A d)$$

$$= \frac{f_{AB}(f_A - v)}{f_A - f_{AB}} = \frac{(v - f_A) + \sqrt{(v - f_A)^2 - 4f_B(v - f_A)}}{2} \tag{9-21}$$

$$f_A = \frac{f_{AB}d}{f_{AB} - s'} \tag{9-22}$$

$$f_B = \frac{-s'd}{f_{AB} - s' - d} \tag{9-23}$$

$$d = \frac{f_B s'}{f_B - s'} = f_A + f_B - \frac{f_A f_B}{f_{AB}} = \frac{(v + f_A) + \sqrt{(v - f_A)^2 - 4f_B(v - f_A)}}{2} \tag{9-24}$$

If a minimum $\Sigma\phi$ is desired where v and f_{AB} are given, one sets

$$d = v - f_{AB} \pm \sqrt{f_{AB}(f_{AB} - v)} \tag{9-25}$$

and

$$s' = f_{AB} \pm \sqrt{f_{AB}(f_{AB} - v)} \tag{9-26}$$

9.2.4. Paraxial Ray Trace – Surface by Surface. It is often more convenient to locate images, determine focal lengths, and so on, by tracing paraxial (first-order) rays through an optical system surface by surface, especially when precise values are required.

Given an optical system with radii r_1, r_2, r_3, ..., r_n separated by distances t'_1, t'_2, t'_3, ..., t'_{n-1} and an object of height h an axial distance l_1 from the vertex of r_1, then rays may be traced through the system by repeated application of the following equations:

$$(u'_i n'_i) = u_{i-1} n_i + y_i (n'_i - n_i)/r_i \tag{9-27}$$

$$y_{i+1} = y_i - t'_i (u'_i n'_i)/n'_i \tag{9-28}$$

where $u'_i = u_{i+1}$ is the slope of the ray after refraction at surface i and y_i is the height that the ray strikes surface i.

Focal lengths, focal points, image positions, and image sizes may be found by the process described in Sec. 9.2.3.1 and given by Eqs. (9-14) through (9-19).

Where the axial intercept of a ray is the only item of interest, the following equations may be used:

$$\frac{n'}{l'} = \frac{n}{l} + \frac{(n' - n)}{r} \tag{9-29}$$

and

$$l_{i+1} = l'_i - t'_i \tag{9-30}$$

More detailed information on ray tracing is provided in Sec. 9.5.

9.3. Limitation of Rays by Stops and Apertures

Apertures, or stops, limit the passage of energy through an optical system.

9.3.1. The Aperture Stop. By following the path of the axial rays in Fig. 9-5, the aperture stop can be determined. Diaphragm #1 is the aperture stop of the system. This limits the size of the axial cone of energy from the object since all the other elements are large enough to accept a bigger cone. The ray through the center of the aperture stop is called the principal or chief ray and is used to locate the pupils. The entrance and exit pupils of the system are the images of the aperture stop formed by all elements in object and image space, respectively. In Fig. 9-5, the entrance pupil lies in the objective lens and the exit pupil to the right of the eye lens. The intersection of the principal ray with the axis locates the pupils, and the diameter of the axial cone at the pupil indicates the pupil diameter.

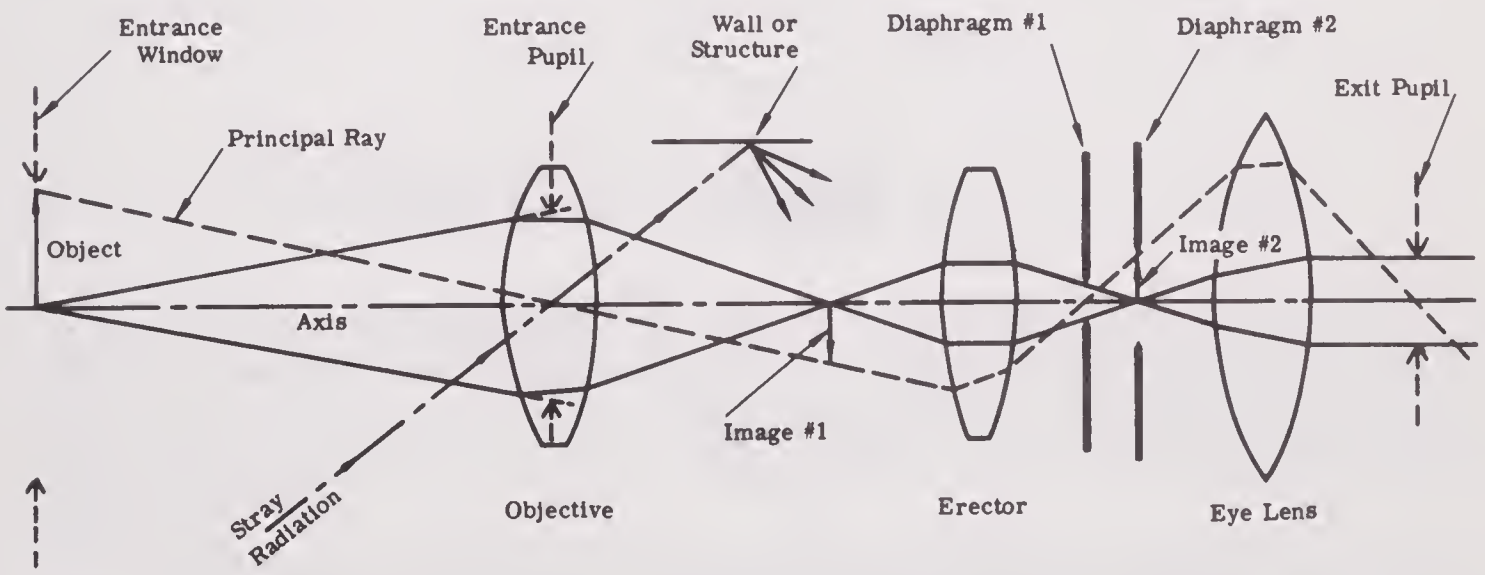


FIG. 9-5. Optical system demonstration entrance and exit pupils and windows, field stop, and glare stops.

9.3.2. The Field Stop. By tracing the path of the principal ray in Fig. 9-5, the field stop can be determined. Diaphragm #2 is the field stop since this diaphragm would prevent a principal ray starting from a point in the object that is further from the axis from passing through the system. The images of the field stop in object and image space are the entrance and exit windows, respectively. In Fig. 9-5, the entrance window is coincident with the object and the exit window is at infinity (which is coincident with the image). The windows of a system do not always coincide with object and image.

The angular field of view of an optical system is the angle that the entrance or exit window subtends from the center of the entrance or exit pupil, respectively.

9.3.3. Vignetting. In Fig. 9-5 the roles played by the various elements of the system are definite and clear cut. In a real system, the diaphragms and lens apertures often play dual roles.

In the system shown in Fig. 9-6(a), the situation is clear for the axial bundle of rays: the aperture stop is the clear aperture of lens A, the entrance pupil is at A and the exit pupil is the image of A formed by lens B. Some distance off axis, however, the cone of energy accepted from point D is limited on its lower edge by the lower rim of lens A and on its upper edge by the upper rim of lens B. This effect is called vignetting, and the appearance of the system aperture when viewed from the object is shown in Fig. 9-6(b). For some object point still farther from the axis, no energy would pass through the system.

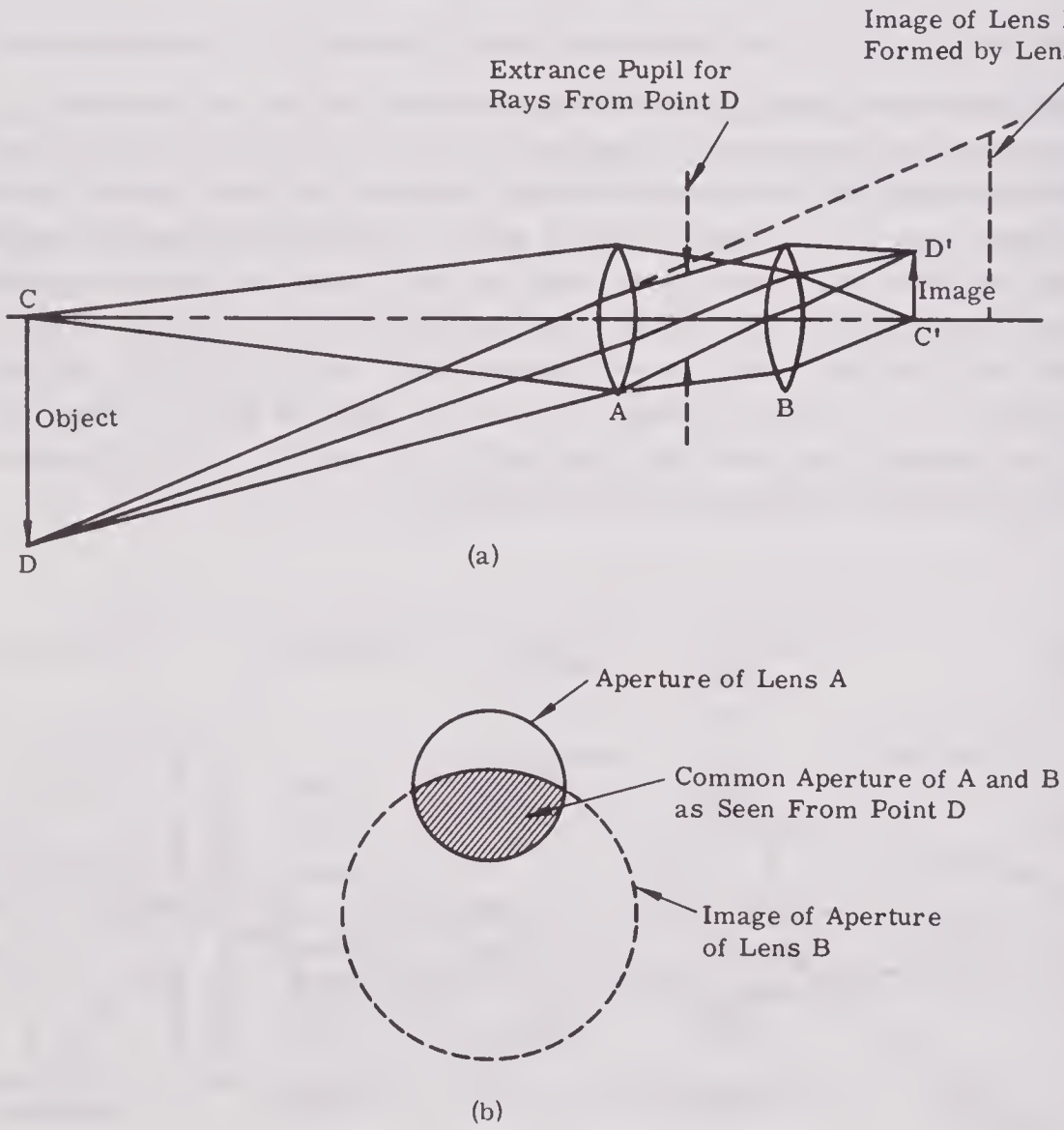


FIG. 9-6. Vignetting. (a) The limitation of the rays from an off-axis point D by the aperture of components A and B. (b) Appearance of the system aperture when viewed from point D.

Thus, for off-axis points, the entrance pupil has become the common area of two circles, one the clear diameter of lens A and the other the diameter of B as imaged by lens A. The dashed lines in the figure indicate the location and size of the image of B, and the arrows indicate the effective entrance pupil, which has a size and position completely different than that for the axial case.

9.3.4. Glare Stops and Baffling. A glare stop is an auxiliary diaphragm located at an image of the aperture stop to block out stray radiation. In Fig. 9-5 the objective diameter could be reduced sufficiently so that it is the aperture stop. Reflected radiation is blocked out by diaphragm #1, acting as a glare stop, if it is an accurate image of the objective, since the stray radiation will appear to be coming from a point outside the objective aperture. Another glare stop could be placed at the exit pupil in this particular system.

In an analogous manner, an auxiliary field stop could be placed at image #1 to further reduce stray radiation.

Baffles are often used to reduce the amount of stray radiation in a system. In Fig. 9-7, if radiation from outside the field of view is a problem, the walls of the system can be baffled as shown in the lower part of the figure. The key to efficient use of baffles is to arrange them so that no part of the detector can "see" a surface which is directly illuminated. The short-dash lines of the figure indicate the stray radiation from the lens. The long-dash lines indicate the limits of the portion of the system visible to the detector, and the solid lines indicate the limits of the illumination which can pass through the lens.

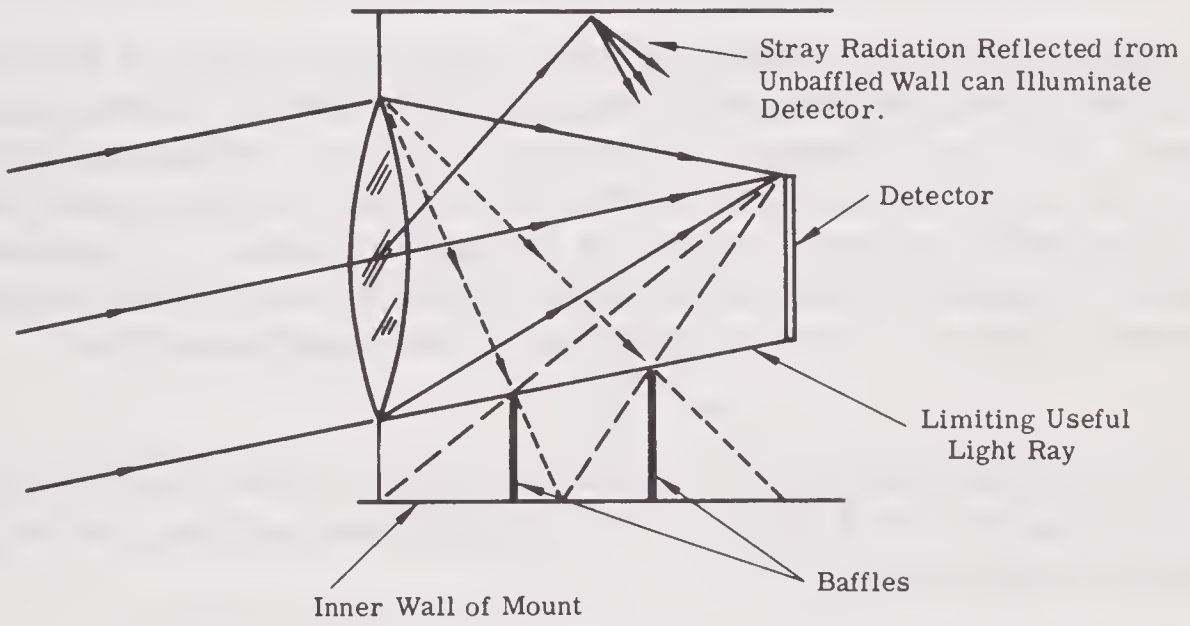


FIG. 9-7. Infrared detector system, showing baffles reducing stray light.

9.3.5. Relative Aperture, Speed, F/No, Numerical Aperture. As noted in Sec. 9.3.1, the irradiance at the image formed by an optical system is limited by the size of the aperture stop. It is in fact determined by the *relative* size of the aperture stop, and thus the illuminating power of an optical system is frequently expressed in terms of relative aperture.

For a system with its object at infinity, the relative aperture is given by the ratio of the effective focal length to the clear aperture (diameter of the entrance pupil). Thus

$$f/\text{no} = \text{e.f.l./clear aperture} \tag{9-31}$$

Another way of expressing this relationship is by the numerical aperture, which is the index of refraction (of the medium in which the image lies) times the sine of the half-angle of the cone of illumination. Thus

$$\text{numerical aperture} = \text{N.A.} = n \sin U \tag{9-32}$$

For aplanatic systems, one free from spherical aberration and coma, numerical aperture and f/no are related by

$$f/\text{no} = \frac{1}{2\text{N.A.}} \tag{9-33}$$

The irradiance on the image is directly proportional to the reciprocal of the square of the f/no, and is given by (see Fig. 9-8)

$$H = \tau\pi N \sin^2 U = \frac{\tau N \pi}{4(f/\text{no})^2} \tag{9-34}$$

where H is the irradiance at the image

τ is the transmission of the system

N is the radiance of the object

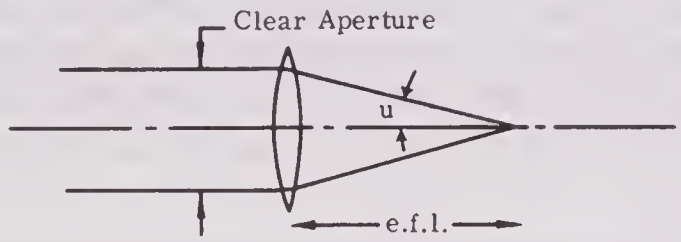


FIG. 9-8. Relationships between f/no = e.f.l./clear aperture − 1/2 N.A. = 1/2 $n \sin u$.

9.3.6. Depth of Field and Depth of Focus. The concept of depth of field assumes that for a given system there exists a blur small enough that it will not affect the system performance. It is then of interest to determine the amount of defocusing which corresponds to this blur size and which can thus be tolerated. In photography the blur size is conventionally expressed in terms of its linear dimension. In infrared work the concept of an angular blur size is more useful. Thus, if a system with clear aperture A will tolerate an angular blur of β radians, one can see from Fig. 9-9 that

$$\delta = \beta D^2/A \tag{9-35}$$

where δ is the distance the object can be shifted from its focused position before it introduces an angular blur of β, and D is the distance from aperture A to the object. (D is assumed large compared to δ.)

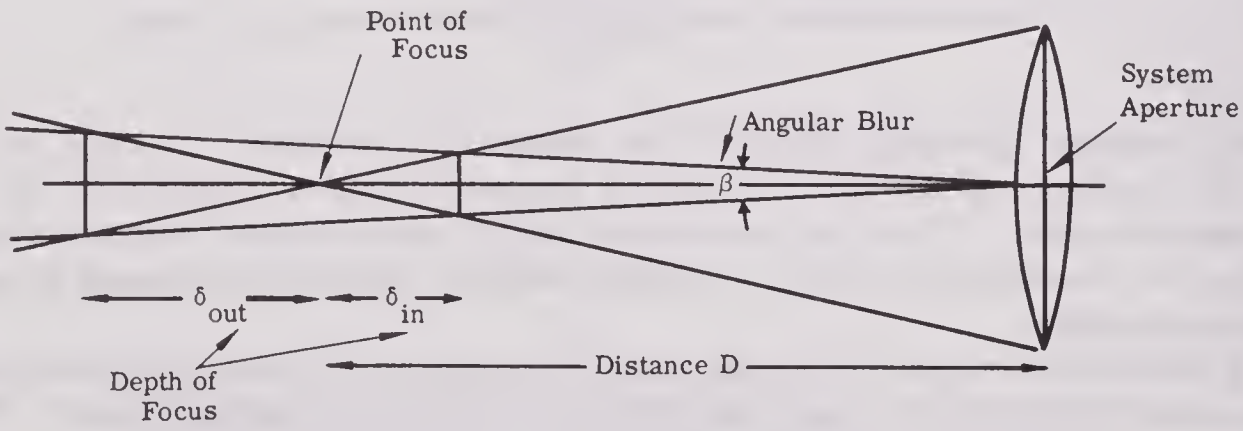


FIG. 9-9. Depth of focus S resulting from a tolerable angular blur.

Similarly,

$$\delta' = \beta D'^2/A = \beta f^2/A = \beta f(f/\text{no}) \tag{9-36}$$

which is applicable when the object is at infinity and the distance (D') from aperture to image is equal to the focal length.

If δ is not small compared with D, then

$$\delta = \beta D^2/(A \pm \beta D) \tag{9-37}$$

and the depth of focus toward the optical system will be smaller than that away from the system.

9.4. Aberrations

Aberrations are departures from perfect imagery in an optical system. A perfect optical system would bring all the rays from a point object to focus at the gaussian image point; because of aberrations some of the rays do not focus at the proper image point.

9.4.1. Aberration Description. An aberration is often expressed as a longitudinal, transverse, or angular aberration. Figure 9-10 uses spherical aberration to show the relationship among the three measures of that aberration.

See Fig. 9-10 for the relationship expressed as follows:

$$LA = \frac{TA}{\tan U'_m} = \frac{AA \cdot f}{\tan U'_m} \tag{9-38}$$

where f is the system focal length (or the distance from the second principal plane to the focal plane if the system is working at finite conjugates), and LA, TA, and AA refer to longitudinal, transverse, and angular aberrations, respectively.

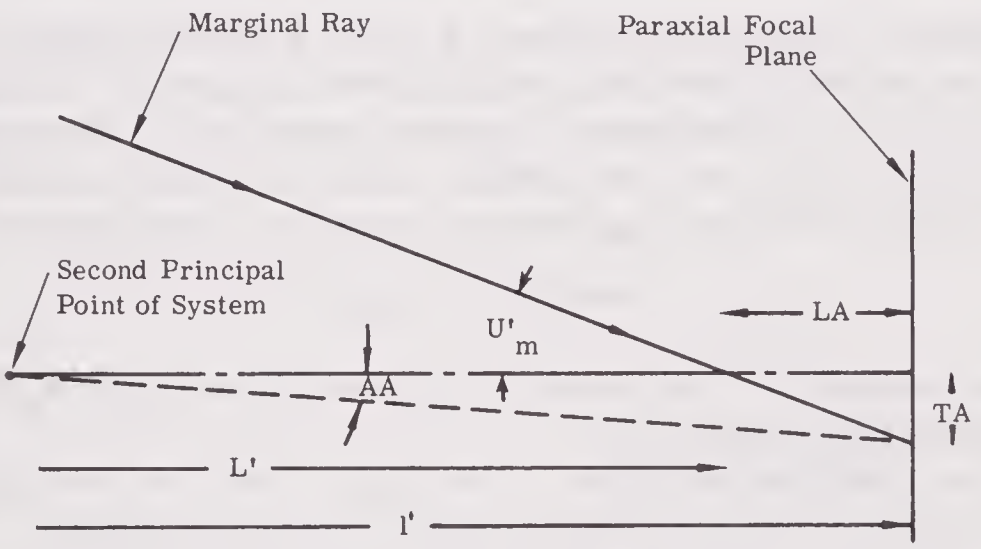


FIG. 9-10. Relationship between longitudinal, transverse, and angular aberration (uncorrected spherical aberration).

Aberrations contributed to a system by a particular element, component, or member are transferred through the system according to the laws of magnification, as outlined in Sec. 9.2.1.2. The angular aberration at the final image is the sum of the angular aberrations of all the contributing elements of the system.

Aberrations are often described as undercorrected or overcorrected. An aberration which is similar in direction or sign to that of a simple positive lens is usually called undercorrected. The sign convention for each aberration is given with the descriptions of the individual aberrations in Sec. 9.4.2.

9.4.2. The Seidel Aberrations. The seven basic aberrations defined below are usually referred to as the Seidel or the primary aberrations.

Mathematically, transverse aberrations can be expressed as a series expansion in terms of the angle U'_m between the marginal ray and the optical axis. The first-order term corresponds to a focusing effect and can be eliminated by a shift of the image reference plane (usually to the position of the paraxial focus). The remaining terms of third, fifth, seventh, and so on, powers of U'_m form a rapidly converging series of which the third-order term is very often predominant. This third-order term is thus a useful approximation of the image aberrations of a system, and the third-order aberrations, as they are called, can be calculated from data derived from a paraxial ray trace.

9.4.2.1. *Spherical Aberration.* Spherical aberration is the variation of focus with aperture, in which a ray through the margin of the lens intersects the axis at a point other than the paraxial focus. A system is undercorrected if the marginal ray comes to a focus before the paraxial focus (see Fig. 9-10). The sign convention for spherical aberration is given by

$$LA' = L' - l' \tag{9-39}$$

where L' is the axial intercept of the marginal ray and l' is the paraxial intercept. The image formed by a system with spherical aberration is a circular blur.

9.4.2.2. *Coma.* Coma is the variation of magnification (or focal length) with aperture. Because of coma, rays passing through the margins of the lens intersect the final image plane at a different height than the principal ray (which passes through the center of the aperture). The upper and lower rim rays of a comatic system intersecting the image plane below the principal ray and the appearance of a typical coma patch are shown in Fig. 9-11. The size of the coma patch is given by

$$\text{coma}_t = H'_{AB} - H'_{pr} \tag{9-40}$$

The correspondence between the position of a ray as it passes through the aperture of a system and its position in the coma patch is shown in Fig. 9-12. As the ray position moves 90° from A to C in the aperture, its position moves 180° in the coma patch. The distance from AE to PR in the coma patch is called tangential coma (coma_t) and the distance from CG to PR is called sagittal coma (coma_s). For third-order coma

$$\text{coma}_t = 3 \cdot \text{coma}_s \tag{9-41}$$

In a typical coma patch 50 to 60 percent of the energy is contained in the pointed end of the patch, between PR and CG.

The Abbé sine condition states that, for an optical system to have perfect imagery in the region near the optical axis (but not at the axis), the ratio $\sin U_1 : \sin U'_k$ must be constant for all rays. The offense against the sine condition (O.S.C.) is a convenient measure of coma in the region of the optical axis and is given by

$$\text{O.S.C.} = \frac{\text{coma}_s}{h} = \frac{\sin U_1}{u_1} \cdot \frac{u'_k}{\sin U'_k} \cdot \frac{l' - l'_{pr}}{(L' - l'_{pr})} - 1 \tag{9-42}$$

where U_1 and u_1 are the slope angles of the marginal and paraxial rays from the axial object point, U'_k and u'_k are the corresponding angles in image space, L' and l' are the axial intercepts of the marginal and paraxial rays, and l'_{pr} is the final intercept of the principal ray (that is, the position of the exit pupil). The quantity $(\sin U_1)/u_1$ is equal to Y/y (the ratio of the ray heights) when the object is at infinity; also, if the exit pupil is at the last surface of the system, $l'_{pr} = 0$.

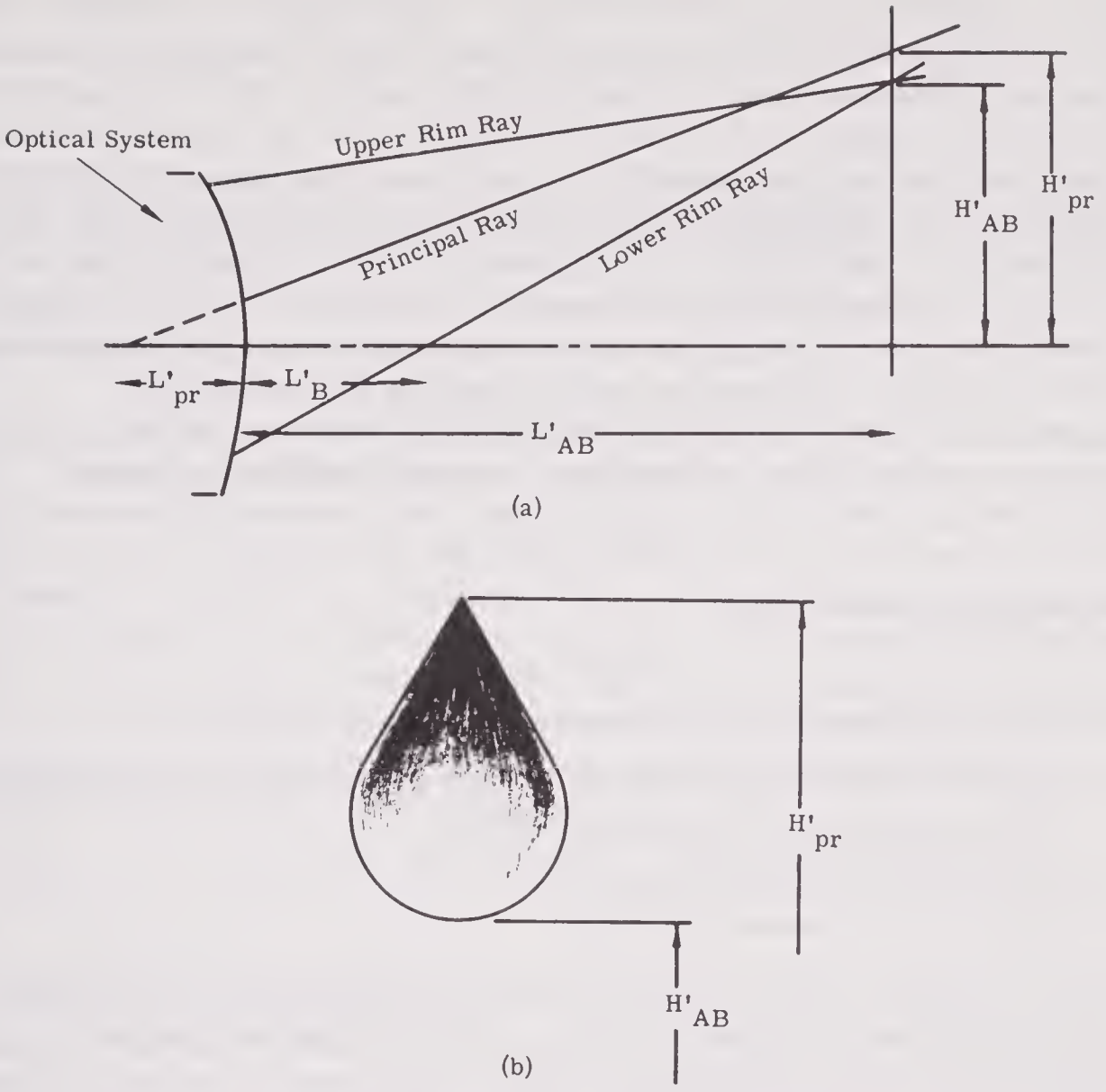


FIG. 9-11. Upper and lower rim rays of comatic system intersecting the image plane below principal ray (*a*); coma patch (*b*).

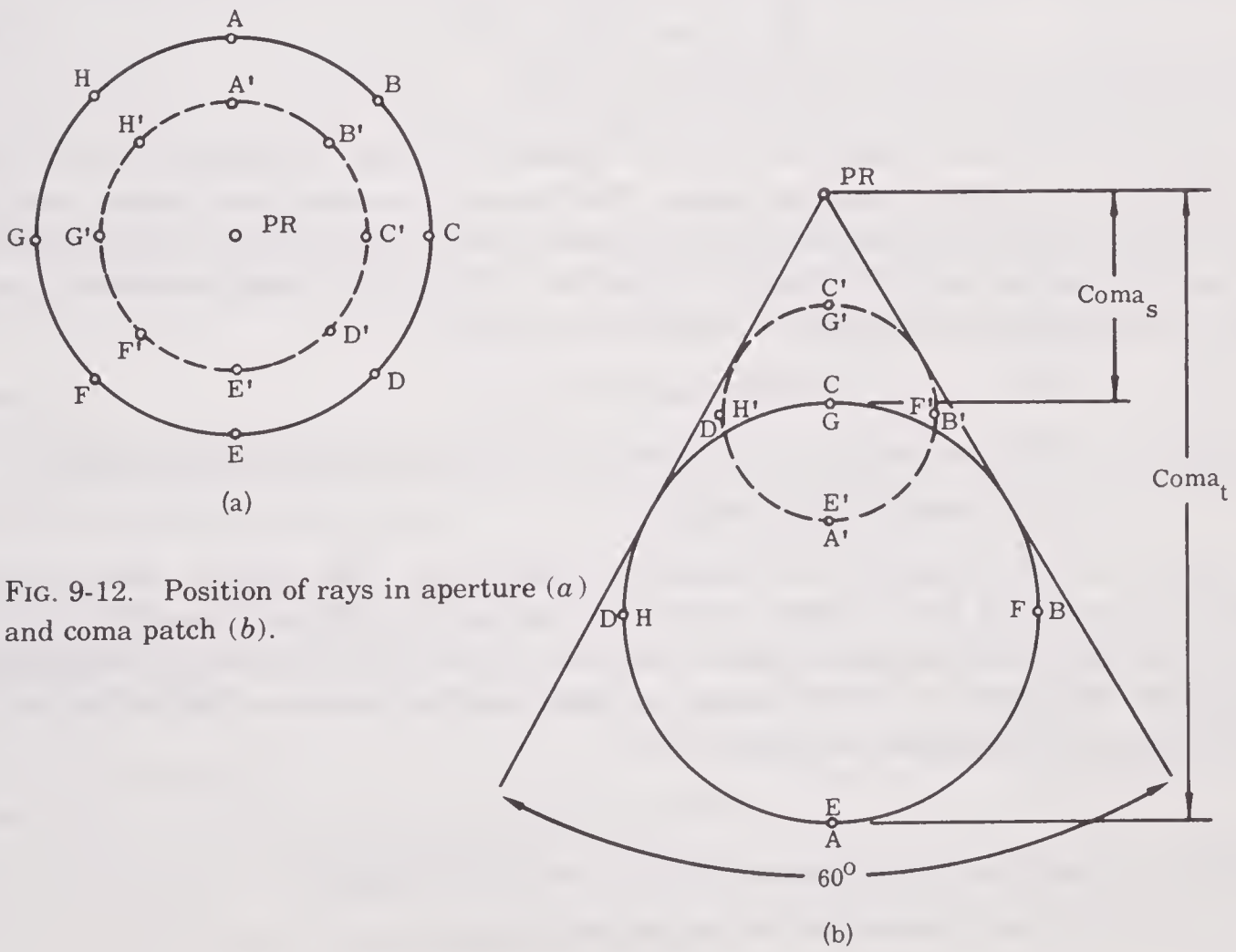


FIG. 9-12. Position of rays in aperture (*a*) and coma patch (*b*).

9.4.2.3. *Astigmatism and Field Curvature.* Astigmatism in an off-axis image of a point is the difference in focus between the fan of rays in the plane of the axis and the object point (the meridional or tangential fan) and the fan of rays perpendicular to this plane (the sagittal or skew fan). In Fig. 9-12(*a*), rays in the plane AA'-$E'E$ of the aperture constitute the tangential fan, and rays in the plane CC'-$G'G$ constitute the sagittal fan. Astigmatism is undercorrected when the tangential fan is brought to a focus before the sagittal fan. The appearance of the astigmatic image is a line at either focus: at the sagittal focus the image is a radial line, which, if extended, would pass through the axis; at the tangential focus the image is a tangential line perpendicular to the radial line. Between the foci the image is diamond-shaped.

The curvature of field is the distance parallel to the axis from the focus of an off-axis image to the axial focal plane. The sagittal curvature of field is given by

$$X_s = L'_{CG} - l' \tag{9-43}$$

and the tangential curvature by

$$X_t = L'_{AE} - l' \tag{9-44}$$

where X is the departure from the plane of the surface of the focus

L' is the axial distance from the last surface of the system to the axial projection of the intersection of the subscript rays

l' is the distance to the paraxial focus

s, CG, t, and AE refer to Fig. 9-12.

An optical system composed of a set of elements of given power and index has a basic field curvature called the Petzval curvature. Although the astigmatism of a system can be changed by bending the elements, the Petzval curvature cannot be changed to any great extent. For regions near the axis, the tangential focus is always three times as far from the Petzval surface as the sagittal focus, satisfying the relationship

$$X_t - X_{\text{ptz}} = 3(X_s - X_{\text{ptz}}) \tag{9-45}$$

where X_{ptz} is the curvature of the Petzval surface.

9.4.2.4. *Distortion.* Distortion is the departure of the image height from that predicted by first-order gaussian optics. The image of a rectangular figure takes on the shape of a pillow or pincushion with concave sides in the presence of distortion that causes an enlargement of the image. Distortion of the opposite sign produces an image with convex sides, like a barrel. Distortion is given by

$$\text{distortion} = H'_{pr} - h' \tag{9-46}$$

where H'_{pr} is the intersection of the principal ray with the paraxial image plane

h' is the paraxial image height.

9.4.2.5. *Axial (Longitudinal) Chromatic Aberration.* The images formed by an optical system may have different sizes and positions for different wavelengths. Because the index of refraction of optical materials varies with wavelength, longitudinal chromatic aberration is the difference in focal position between images formed by two different wavelengths, and is given by

$$L\text{ch}A = l'_v - l'_r \tag{9-47}$$

where l'_v is the final image distance for the shorter wavelength

l'_r the image distance for the longer wavelength.

9.4.2.6. *Off-Axis (Transverse) Chromatic Aberration.* Off-axis aberration results in a difference in image size due to wavelength variation and is given by

$$T\text{ch}A = H'_v - H'_r \tag{9-48}$$

where H'_v = paraxial image plane intersection of principal rays of short wavelengths

H'_r = paraxial image plane intersection of principal rays of long wavelengths

Transverse chromatic, or lateral color can be considered as a chromatic difference of magnification (C.D.M.) and expressed as:

$$\text{C.D.M.} = (H'_v - H'_r)/h' \tag{9-49}$$

9.4.3. Correction of Aberrations. In practical optical systems, aberrations are usually corrected by balancing the undercorrection of one element against the overcorrection of another. Because an optical system should have a given power of focal length, it is necessary to combine elements whose $\Sigma y\phi$ (see Sec. 9.2.3.1) equals the desired power, but whose summed aberrations equal zero. In chromatic aberration in a doublet, a positive element with a certain chromatic aberration contribution per unit of power is combined with a negative element with a relatively higher chromatic contribution per unit of power, so that the chromatic contributions are equal and opposite and cancel while leaving a residue of power.

Chromatic aberrations are corrected by proper choice of materials, element powers, and spacings. The Petzval curvature is corrected by choice of material and element powers. The preceding aberrations are ordinarily corrected in the designer's initial layout of the powers and spacings of the optical elements to be used in the system, at the same time that the system power and working distance are arranged. Spherical aberration, coma, astigmatism, and distortion can be controlled by proper shaping or bending of the elements of the system. Aspheric surfaces (that is, surfaces of revolution which are not spheres) may also be used to correct the monochromatic aberrations.

9.4.4. Variations of Aberrations with Aperture and Image Size. The amount of a primary aberration in the image is a function of the semiaperture (y) of the system and the height (h) of the image. Table 9-1, based on third-order aberrations, is useful in estimating the effect of a change in aperture or field coverage (image size) on the performance of a system.

TABLE 9-1. VARIATION OF ABERRATIONS WITH SIZE OF APERTURE, FIELD ANGLE, AND IMAGE SIZE, FOR SYSTEM WITH PURE THIRD-ORDER ABERRATIONS

	Size of Aperture	*Field Angle*	*Image Size*
Aberration	(y)	(u)	(h)
Longitudinal spherical aberration	y^2	—	—
Transverse spherical aberration	y^3	—	—
Coma	y^2	u	h
Astigmatism	—	u^2	h^2
Length of astigmatic focal lines	y	u^2	h^2
Petzval curvature	—	u^2	h^2
Distortion	—	u^3	h^3
Percentage distortion	—	u^2	h^2
Longitudinal axial chromatic aberration	—	—	—
Transverse axial chromatic aberration	y	—	—
Lateral chromatic	—	u	h
Chromatic difference of magnification (C.D.M.)	—	—	—

9.4.5. Zonal and Residual Aberrations. When an aberration is fully corrected for a certain region of the aperture or portion of the field, there usually remain aberrations for rays passing through other parts of the aperture or for smaller or larger field angles.

9.4.6. Chromatic Variation of Aberrations. Because the index of optical materials changes with wavelength, the monochromatic aberrations (spherical, coma, astigmatism, Petzval curvature, and distortion) will also vary with the wavelength of radiation passing through an optical system. Chromatic variation of spherical aberration (spherochromatism) is most common. Ordinary spherochromatism causes the spherical aberration in the shorter wavelength to be more overcorrected than that of the longer wavelengths.

9.4.7. Graphical Representation of Aberrations. Spherical and chromatic aberration can be presented as longitudinal or transverse aberrations, the former being plotted against the entering ray height (or angle) and the latter against the slope angle of the emergent ray. The spherical aberration of a single element with both types of graphical presentation is shown in Fig. 9-13.

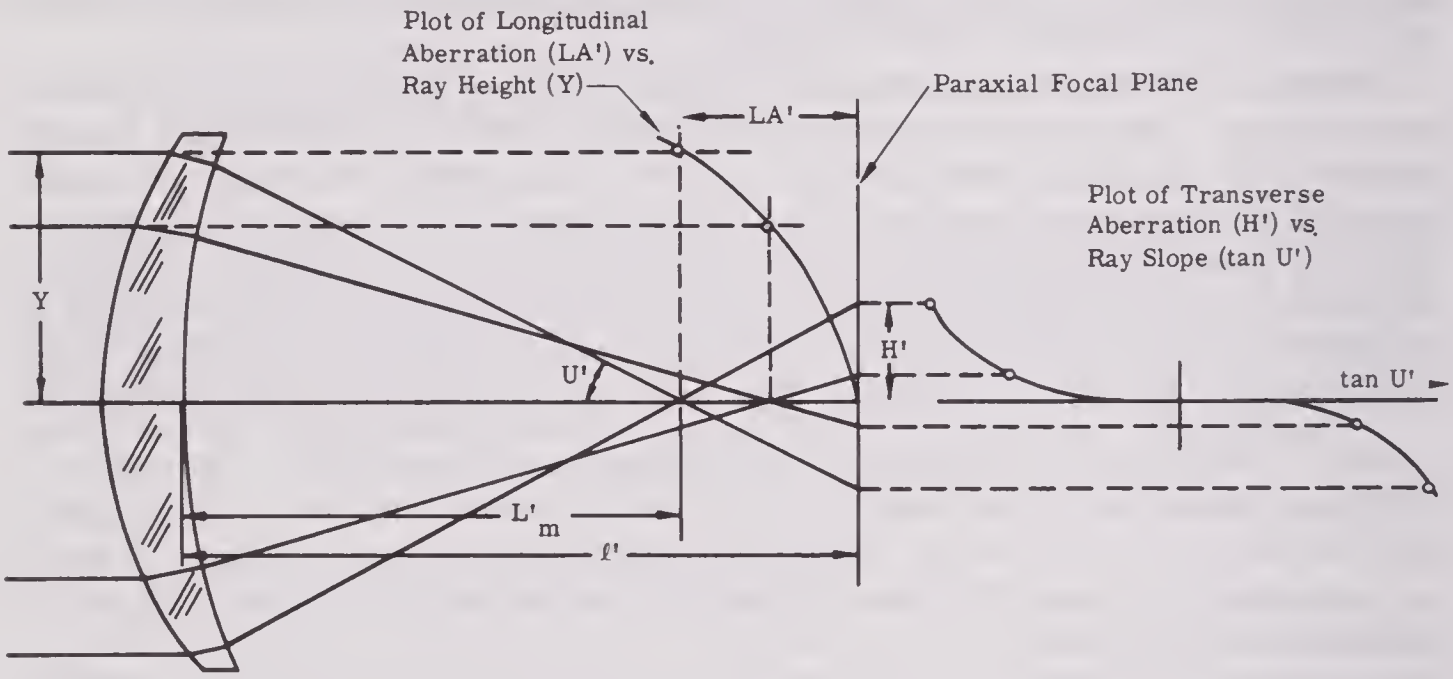


FIG. 9-13. Undercorrected spherical aberration. The longitudinal spherical aberration is plotted against the entering ray height Y. The transverse spherical is plotted against the tangent of U', the final angle the light ray makes with the axis.

The plot of H' *vs.* tan U' is called a *rim ray curve* and is useful because it indicates directly the size of the blur in the image caused by the aberration; and the effect of refocusing (or shifting the reference plane) can be readily determined by rotating the X or tan U' axis.

Curvature of field is ordinarily represented by plotting the longitudinal difference of focus between axial and oblique rays against either the image height or the field angle, as shown in Fig. 9-14.

The rim ray curve can also be used to represent the aberrations of oblique fans of meridional rays, as shown in Fig. 9-15. The shape of the rim ray curve is indicative of the aberration present, as shown in Fig. 9-16 for common aberrations.

Meridional aberrations can be expressed in terms of X and Y. Skew or sagittal fans have aberrations in the Z direction also, and the representation of the aberrations of skew rays is more complex. Two methods are common. One is to plot the Y-Z

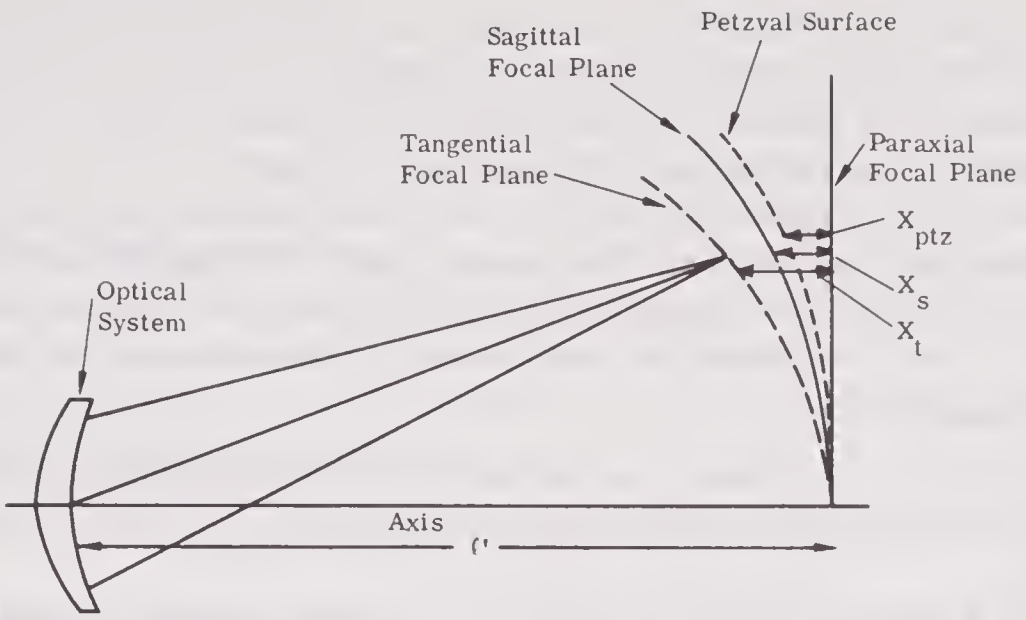


FIG. 9-14. Relationships between sagittal and tangential focal surfaces and Petzval surface for lens with undercorrected astigmatism.

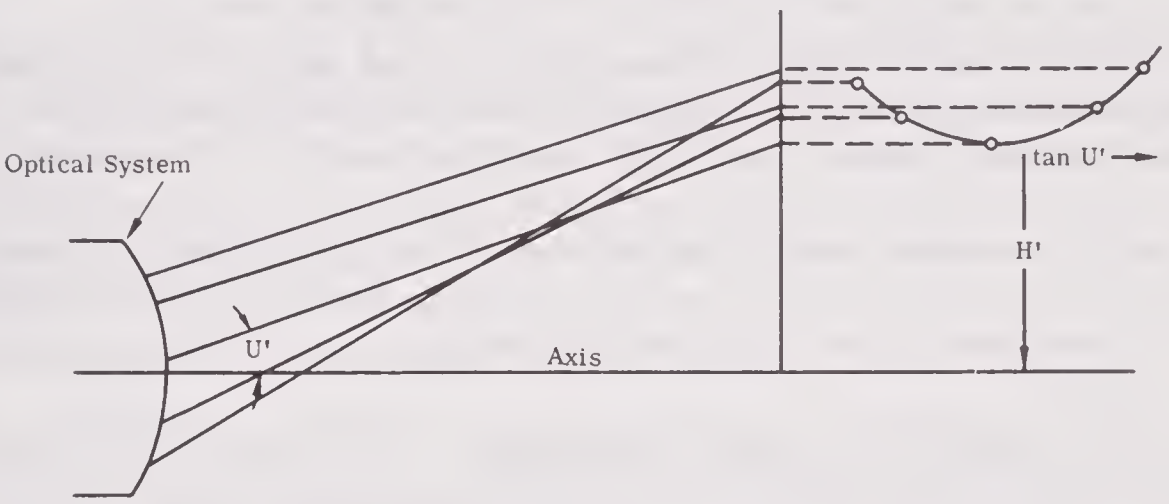


FIG. 9-15. Construction of off-axis rim ray curve to represent aberrations of oblique fans of meridional rays.

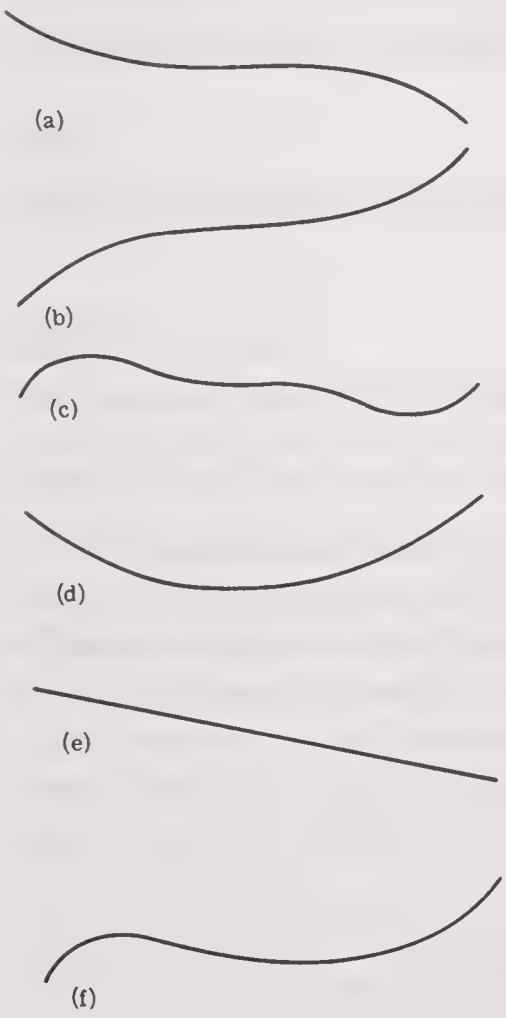


FIG. 9-16. Appearance of rim ray curve in presence of typical aberrations: (*a*) spherical aberration, third-order undercorrected, (*b*) spherical aberration, third-order overcorrected, (*c*) zonal spherical aberration, (*d*) third-order coma, (*e*) defocusing or field curvature, (*f*) rim ray curve for a typical oblique fan showing coma, field curvature, and oblique spherical.

coordinates of the ray intersections with the image plane, connecting the points made by rays which lie along a radial line in the aperture. If a large number of points is plotted, the density of points is representative of the flux density in the image, and the flux density in the image and the spot diagram become a "picture" of the image (provided that diffraction can be neglected). In another method, the Y and Z departures are plotted separately against the radial distance from ray to the center of the aperture. The rim ray curve described above is this type of "clock plot" for the meridional fan. A clock plot is useful in aberration analysis and in interpolating for spot diagrams.

9.5. Ray Tracing

Ray tracing is used to calculate the path of a ray of light through an optical system. It is ordinarily done either to evaluate an optical system or as a step in the process of optical design.

Ray tracing is based on Snell's law (Eq. 9-1) and can be done by using geometry to calculate the ray paths between surfaces and applying Snell's law at each surface. Specific equations for this purpose are given in Sec. 9.5.3. A crude form of ray tracing may be carried out with drawing tools; for the technique see Sec. 9.5.5.

9.5.1. Ray-Tracing Precision. Depending upon the scale of the calculation, the precision required of dimensional numbers is to four, five, or six decimal places. Angles and trigonometrical functions should have six-figure accuracy. Trigonometric tables given in terms of radians are preferred for ray tracing with a desk calculator. Most trigonometric ray-tracing results are referred to the paraxial focus or focal plane. The paraxial ray trace is carried out with equations derived from the trigonometric ray-tracing equations by setting the sine and tangent equal to the angle, and cosines equal to unity.

9.5.2. Determination of Specific Aberrations. The following subsections outline the methods of obtaining numerical values for the basic aberrations.

9.5.2.1. *Spherical Aberration.* A paraxial ray and a marginal ray (a trigonometric ray through the rim of the entrance pupil) are traced starting at the axial point of the object. The final axial intercepts of the rays are to be determined. Longitudinal spherical aberration is given by

$$LA' = L' - l' \tag{9-50}$$

Transverse spherical aberration is the height at which the ray strikes the paraxial image plane and can be found from

$$TA' = LA' \tan U' \tag{9-51}$$

Zonal spherical aberration is found by tracing a ray at a lesser height than the marginal ray (usually at 0.707 of the marginal height) and substituting the final zonal-ray data in Eqs. (9-50) and (9-51) in place of the marginal-ray data.

9.5.2.2. *Coma.* Tangential coma is evaluated by tracing three rays of a meridional fan from an off-axis object point through the system. The principal ray (pr) passes through the center of the entrance pupil, and the upper and lower rim rays (A and B) are traced through the upper and lower edges of the pupil. The intersection of ray A with ray B is determined, and the distance from the intersection to the axis (H'_{AB}) is compared to the height of the intersection of the principal ray with a plane through the AB intersection (H'_{pr}). The axial distance from the last surface to the AB intersection can be found from

$$L'_{AB} = \frac{L'_A \tan U'_A - L'_B \tan U'_B}{\tan U'_A - \tan U'_B} \tag{9-52}$$

and the intersection heights from

$$H'_{AB} = (L'_A - L'_{AB}) \tan U'_A \tag{9-53}$$

$$H'_{pr} = (L'_{pr} - L'_{AB}) \tan U'_{pr} \tag{9-54}$$

Then tangential coma is given by

$$\text{coma}_T = H'_{AB} - H'_{pr} \tag{9-55}$$

The offense against the sine condition (O.S.C.) can be determined from the same ray trace used to find spherical aberration:

$$\text{O.S.C.} = \frac{\sin U_1}{u_1} \cdot \frac{u'}{\sin U'} \cdot \frac{(l' - l'_{pr})}{(L' - l'_{pr})} - 1 \tag{9-56}$$

For regions close to the optical axis,

$$\text{coma}_t = 3 \text{ coma}_s = 3H'\text{O.S.C.} \tag{9-57}$$

9.5.2.3. *Astigmatism and Field Curvature.* Curvature of field for small apertures is found by tracing the equivalent of a paraxial ray near a principal ray instead of near the optical axis. The calculation is carried out using Coddington's equations (see Sec. 9.5.3).

The extended Petzval surface may be found from the x_s and x_t resulting from a Coddington's trace by

$$x_{\text{ptz}} = 1.5x_s - 0.5x_t \tag{9-58}$$

9.5.2.4. *Distortion.* Distortion is found by tracing a trigonometric principal ray and comparing the height of its intersection with the paraxial focal plane to the image height predicted by gaussian optics. The height for the principal ray is found from

$$H'_{pr} = (L'_{pr} - l') \tan U'_{pr} \tag{9-59}$$

For objects at infinity the gaussian height is given by

$$h' = -f \tan U_{pr} \tag{9-60}$$

and for an object at a finite distance by

$$h' = hnu/n'u' \tag{9-61}$$

Distortion is then

$$dist = H'_{pr} - h' \tag{9-62}$$

Distortion is often specified as a percentage of the image height.

9.5.2.5. *Longitudinal Chromatic Aberration.* This aberration is found by locating the image position for the different wavelengths of interest by using the appropriate values for the index of refraction in the calculation. The calculation may be paraxial or trigonometric. For paraxial rays

$$L\text{ch}A = l'_v - l'_r \tag{9-63}$$

where v and r refer to the short and long wavelengths, respectively. If marginal rays are traced in various wavelengths, it is useful to make a plot of L' *vs.* Y, or H' *vs.* tan U' (see Sec. 9.4.7) for all wavelengths on the same graph, because this will also indicate the spherochromatism.

9.5.2.6. *Lateral Color, or Transverse Chromatic Aberration.* Lateral color is found by tracing principal rays for the long and short wavelengths. The lateral color is the difference between the paraxial focal-plane intersection heights of the two principal rays and is given by

$$T\text{ch}A = H'_v - H'_r \tag{9-64}$$

where H'_v and H'_r are determined from Eq. (9-59).

9.5.3. Ray-Tracing Equations.

9.5.3.1. *For Desk Calculators. Paraxial ray-tracing equations* (see Fig. 9-17):

Opening:

$$n_1 u_1 = n_1 y_1 / l_1 \tag{9-65}$$

Iterative:

$$n'_i u'_i = n_i u_i + y_i(n'_i - n_i)/r_i \tag{9-66}$$

$$y_{i+1} = y_i - t'_i n'_i u'_i / n'_i \tag{9-67}$$

Closing:

$$l'_i = y_i n'_i / n'_i u'_i \tag{9-68}$$

If $u_1 = 0$, then

$$\text{e.f.l.} = y_1 n'_k / n'_k u'_k \tag{9-69}$$

$$\text{b.f.l.} = y_k n'_k / n'_k u'_k \tag{9-70}$$

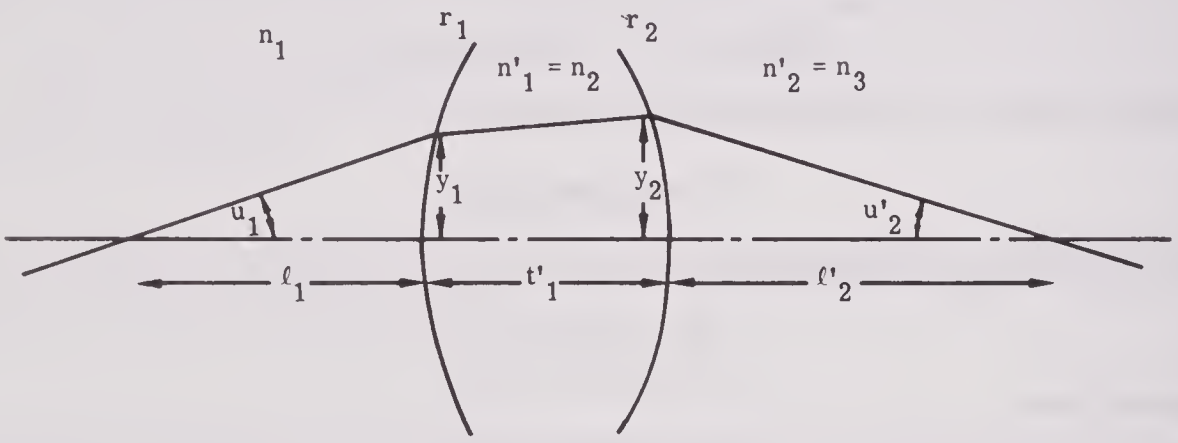


FIG. 9-17. Quantities used in paraxial ray tracing.

For speed, the quantity nu is carried as an entity. If the values of the angles of incidence and refraction are required for third-order aberration calculations, they can be found from

$$i = y/r - u \tag{9-71}$$

$$i' = ni/n' \tag{9-72}$$

Meridional ray-tracing equations for finite radii (see Fig. 9-18):

Opening:

$$\overline{CA} = (L - R) \sin u \tag{9-73}$$

or if $U = 0$, $CA = y$.

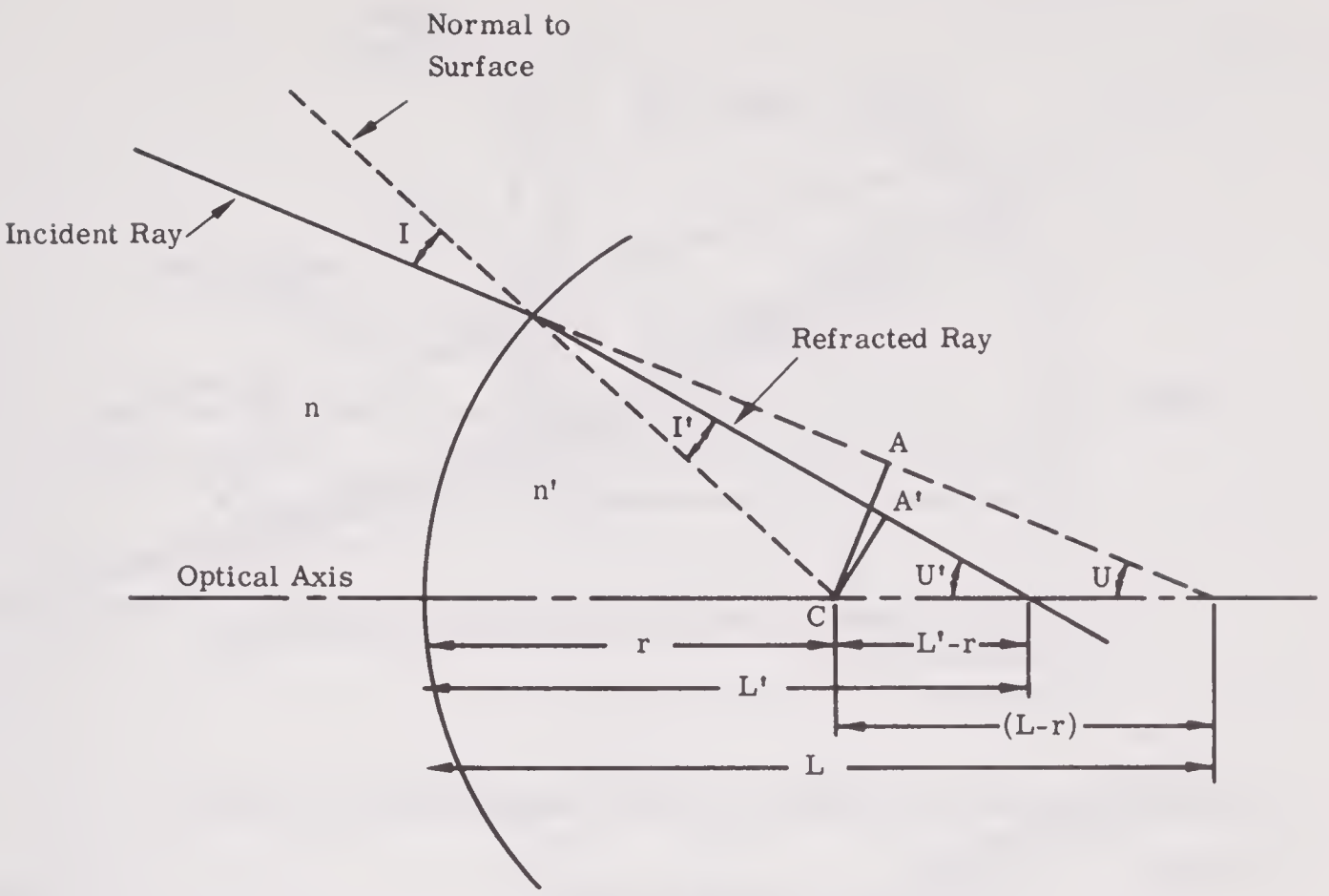


FIG. 9-18. Quantities used in meridional ray-trace equations for spherical surfaces.

Iterative:

$$\sin I_i = CA_i/r_i \tag{9-74}$$

$$\sin I'_i = \frac{n_i}{n'_i} \sin I_1 \tag{9-75}$$

$$U'_i = U_i + I_i - I'_i \tag{9-76}$$

$$U_{i+1} = U'_i$$

$$CA_{i+1} = (r_i - r_{i+1} - t'_i) \sin U'_i + \frac{n_i CA_i}{n'_i} \tag{9-77}$$

Closing:

$$L'_i = r_i + \frac{n_i CA_1}{n'_i \sin U'_i} \tag{9-78}$$

Miscellaneous:

Coordinates of ray intersection with surface:

$$Y = r \sin (U + I) \tag{9-79}$$

$$X = 2r \sin^2 \tfrac{1}{2} (U + I) \tag{9-80}$$

Distance between surfaces along ray:

$$D_{i \text{ to } i+1} = (t'_i + x_{i+1} - x_i) \sec U'_i \tag{9-81}$$

or

$$D_{i \text{ to } i+1} = r_i \cos I'_i - r_{i+1} \cos I_{i+1} - (r_i - r_{i+1} - t'_i) \cos U'_i \tag{9-82}$$

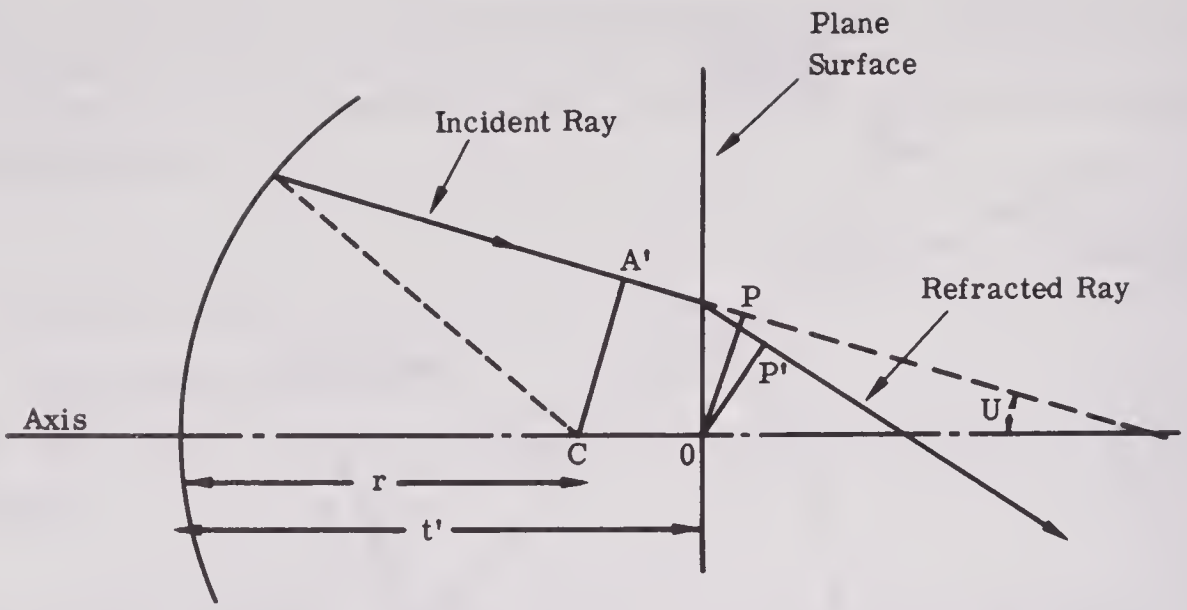


FIG. 9-19. Quantities used in meridional ray-trace equations for plane surfaces.

Meridional ray-tracing equations for plane surfaces (see Fig. 9-19):

Opening:

$$OP_1 = L_1 \sin U_1 \tag{9-83}$$

or if $U_1 = 0$, $OP_1 = Y_1$.

Transfer from preceding radius:

$$OP_i = \frac{n_{i-1} CA_{i-1}}{n'_{i-1}} + (r_{i-1} - t'_{i-1}) \sin U'_{i-1} \tag{9-84}$$

$$U_i = U'_{i-1} \tag{9-85}$$

Iterative:

$$\sin U'_i = \frac{n_i}{n'_i} \sin U_i \tag{9-86}$$

$$OP'_i = \frac{\cos U'_i}{\cos U_i} OP_i \tag{9-87}$$

$$OP_{i+1} = OP'_i - t'_i \sin U'_i \tag{9-88}$$

$$CA_{i+1} = OP'_i + (-r_{i+1} - t'_i) \sin U'_i \tag{9-89}$$

Closing:

$$L'_i = OP'_i / \sin U'_i \tag{9-90}$$

Miscellaneous:

$$Y_i = OP_i / \cos U_i \tag{9-91}$$

$$D_{i \text{ to } i+1} = r_i \cos I'_i - OP_{i+1} \tan U'_i - (r_i - t'_i) \cos U'_i \tag{9-92}$$

$$D_{i \text{ to } i+1} = -(-r_{i+1} - t'_i) \cos U'_i + OP'_i \tan U'_i - r_{i+1} \cos I_{i+1} \tag{9-93}$$

Coddington's equations (for tracing close sagittal and tangential rays about a principal ray for determination of astigmatism and field curvature) (see Fig. 9-20): A principal ray is traced using Eqs. (9-78) through (9-93) (as applicable).

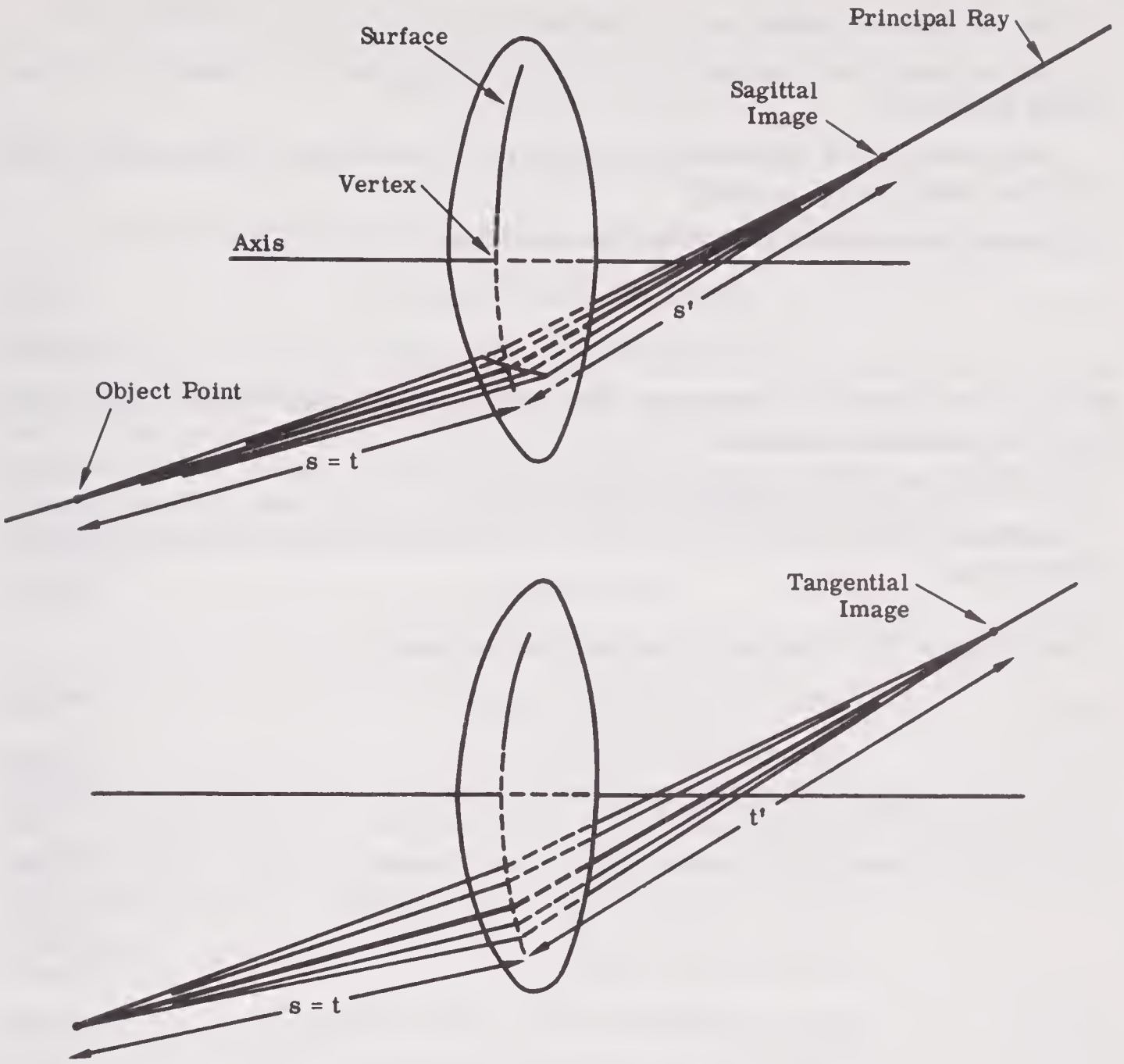


FIG. 9-20. Quantities used in Coddington's equations.

The "oblique power," ϕ, of each surface is calculated from

$$\phi = (N' \cos I'_p - N \cos I_p)/r \tag{9-94}$$

The sagittal rays are traced by repeated application of

$$\frac{n'}{s'} = \frac{n}{s} + \phi \tag{9-95}$$

and

$$s_{i+1} = s'_i - D_{i \text{ to } i+1} \tag{9-96}$$

The tangential rays are traced by repeated application of

$$\frac{n' \cos^2 I'_p}{t'} = \frac{n \cos^2 I_p}{t} + \phi \tag{9-97}$$

and

$$t_{i+1} = t'_i - D_{i \text{ to } i+1} \tag{9-98}$$

where s and t are the distances along the principal ray from surface to focus.

Under the usual sign convention the ray trace is started with $s_1 = t_1$ equal to a negative value if the object is to the left of the first surface.

For an object at infinity, $s_1 = t_1 =$ infinity.

For a finite object distance $s_1 = t_1 = (L_1 - X_{pr_1})$ sec U_{pr_1}, where X_{pr} is found from Eq. (9-80).

The quantity D is the distance along the ray from surface to surface and is found by Eqs. (9-82), (9-92), or (9-93).

Closing: The final curvature of field is found from

$$X'_s = s'_k \cos U'_{p_k} + X_{p_k} - l'_k \tag{9-99}$$

$$X'_t = t'_k \cos U'_{p_k} + X_{p_k} - l'_k \tag{9-100}$$

9.5.3.2. *For Electronic Computers.* The paraxial equations in Sec. 9.5.3.1 are suitable for electronic computers.

Meridional ray-tracing equations (see Fig. 9-21):

Opening: Q is the perpendicular to the ray from the vertex (axial intersection) of the surface; thus

$$Q_1 = L_1 \sin U_1 \tag{9-101}$$

Iterative: c is the reciprocal of the radius of curvature.

$$\sin I = Qc - \sin U \tag{9-102}$$

$$\cos I = \sqrt{1 - \sin^2 I} \tag{9-103}$$

$$\sin (U+I) = \cos U \cos I + \sin U \sin I \tag{9-104}$$

$$\cos (U+I) = \cos U \cos I - \sin U \sin I \tag{9-105}$$

$$\sin I' = (n/n') \sin I \tag{9-106}$$

$$\cos I' = \sqrt{1 - \sin^2 I'} \tag{9-107}$$

$$\sin U' = \sin (U+I) \cos I' - \cos (U+I) \sin I' \tag{9-108}$$

$$\cos U' = \cos (U+I) \cos I' + \cos (U+I) \sin I' \tag{9-109}$$

$$Q' = Q \frac{(\cos U' + \cos I')}{(\cos U + \cos I)} \tag{9-110}$$

$$Q_{i+1} = Q'_i - t'_i \sin U'_i \tag{9-111}$$

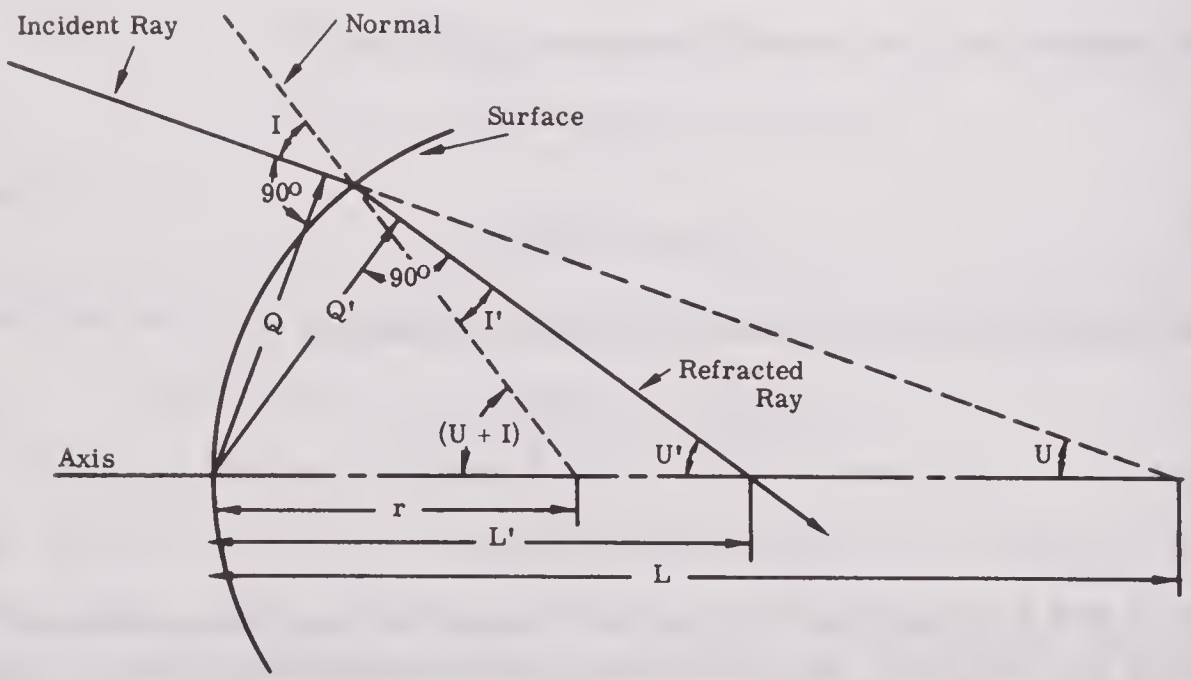


FIG. 9-21. Quantities used in computing formulas for electronic computers.

The intersection height of the ray with the surface may be found from

$$Y = Q' \frac{[1 + \cos (U+I)]}{(\cos U' + \cos I')} \tag{9-112}$$

Closing:

$$L'_i = Q'_i / \sin U'_i \tag{9-113}$$

To find the height the ray strikes a plane a distance l' from the surface i, use

$$H' = (Q'_i - l'_i \sin U'_i)/\cos U'_i \tag{9-114}$$

Coddington's equations (*close skew and close meridional rays*): For electronic computers Coddington's equations have been rewritten in a form which does not contain the quantities s or t directly. A principal meridional ray is traced by the equations of Sec. 9.5.3.2. concurrently with the Coddington trace; the ordinary ray-tracing quantities used in the following refer to data of this principal ray.

Opening:

$$P_s = n_1 y_s \left(\frac{1}{s_1}\right) \tag{9-115}$$

$$P_t = n_1 y_t \cos^2 I_1 \left(\frac{1}{t_1}\right) \tag{9-116}$$

where y is an arbitrarily chosen ray height from the principal ray, analogous to the y of the paraxial ray trace, and s_1 and t_1 are the distances from the object to the first surface along the principal ray, and P_s and P_t are the tangential and sagittal representations for the principal ray.

Iterative:

$$\phi_i = c_i(n'_i \cos I'_i - n_i \cos I_i) \tag{9-117}$$

$$P'_{si} = P'_{s(i-1)} + y_{si}\phi_i \tag{9-118}$$

$$P'_{ti} = P'_{t(i-1)} + y_{ti}\phi_i \tag{9-119}$$

$$x_i = \frac{Q_i \sin (U_i + I_i)}{(\cos U_i + \cos I_i)} \tag{9-120}$$

$$D_{i \text{ to } i+1} = \frac{t_{i \text{ to } i+1} + x_{i+1} - x_i}{\cos U'_i} \tag{9-121}$$

$$y_{s(i+1)} = y_{si} - P'_{si} D_{i \text{ to } i+1}/n_i \tag{9-122}$$

$$y_{t(i+1)} = \frac{\cos^2 I'_i}{\cos^2 I_{(i+1)}} \frac{(y_{ti} - P'_{ti} D_{i \text{ to } i+1})}{n'_i \cos^2 I'_i} \tag{9-123}$$

Closing:

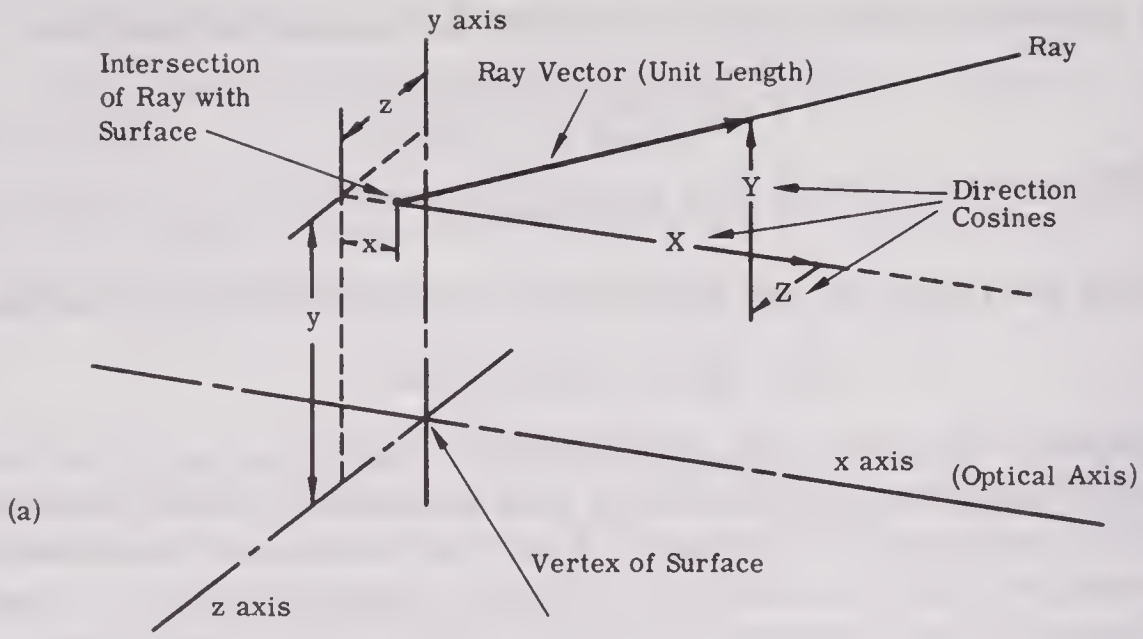
$$s'_k = n'_k y_{sk}/P'_{sk} \tag{9-124}$$

$$t'_k = n'_k y_{tk} \cos^2 I'_k/P'_{tk} \tag{9-125}$$

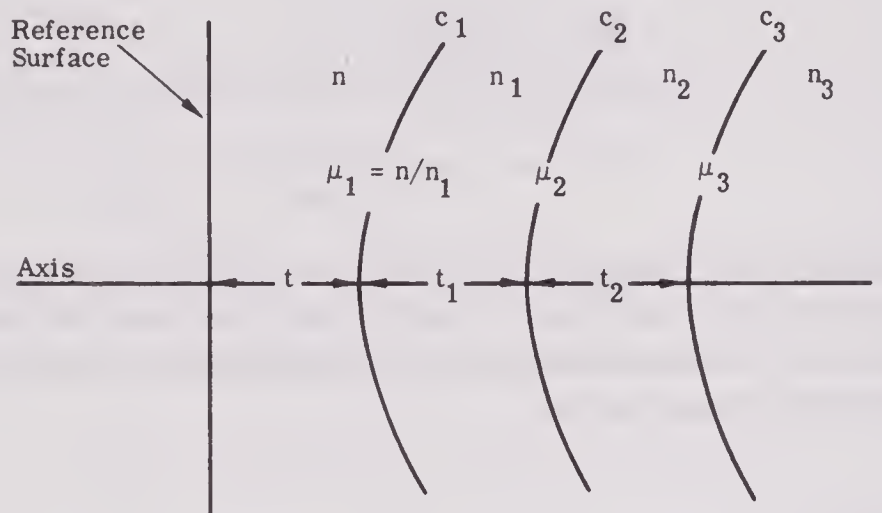
The curvature of field from a surface a distance l' from the final surface may be found from

$$X'_s = s'_k \cos U'_k + X_k - l' \tag{9-126}$$

$$X'_t = t'_k \cos U'_k + X_k - l' \tag{9-127}$$



(a)



(b)

FIG. 9-22. Direction cosines and surfaces intercept coordinates (*a*) and subscript convention (*b*) used in skew ray-trace formulas.

Skew ray trace through a spherical surface (Fig. 9-22*a*): The general ray is defined by its direction cosines X, Y, Z and the coordinates of its intersection with the surface x, y, z. The calculation is usually started with a dummy surface ($c = 0$ and $n = n'$) placed at the entrance pupil for convenience; thus the opening data are the ray vectors **X**, **Y**, **Z** and the ray coordinates in this surface. The calculation is closed by calculating the ray coordinates x, y, z in a final reference surface which is usually the final image plane (or surface). L is the distance along the ray from surface i to surface $(i + 1)$.

The following equations are applied surface by surface:

$$e = tX - xX - yY - zZ \tag{9-128}$$

$$M_x = x + eX - t \tag{9-129}$$

$$M_1^2 = x^2 + y^2 + z^2 - e^2 + t^2 = 2tx \tag{9-130}$$

$$E_1 = \sqrt{x^2 - c_1(c_1 M_1^2 - 2M_x)} \tag{9-131}$$

$$L = e + (c_1 M_1^2 - 2M_x)/(X + E_1) \tag{9-132}$$

$$x_1 = x + LX - t \tag{9-133}$$

$$y_i = y + LY \tag{9-134}$$

$$z_1 = z + LZ \tag{9-135}$$

$$\mu_1 = n/n_1 \tag{9-136}$$

$$E'_1 = \sqrt{1 - \mu_1^2 (1 - E_1^2)} \tag{9-137}$$

$$g_1 = E'_1 - \mu_1 E_1 \tag{9-138}$$

$$X_1 = \mu_1 X - g_1 c_1 x_1 + g_1 \tag{9-139}$$

$$Y_1 = \mu_1 Y - g_1 c_1 y_1 \tag{9-140}$$

$$Z_1 = \mu_1 Z - g_1 c_1 z_1 \tag{9-141}$$

The ray "height" at the surface may be found from

$$s_1 = \sqrt{y_1^2 + z_1^2} \tag{9-142}$$

Unsubscripted quantities refer to surface i and those with the subscript 1 refer to surface $(i+1)$. The axial spacing t is from i to $(i+1)$. n is n_{i+1} and n_1 is n'_{i+1}. E_1 and E'_i are the cosines of the angles of incidence and refraction. See Fig. 9-22(b).

Skew ray trace through an aspheric surface: As above, the general ray is defined by its direction cosines X, Y, Z and its intersection coordinates x, y, z. The difficulty in tracing through an aspheric lies in determining the intersection of the ray and the aspheric surface. This is accomplished by successive approximations, the approximation process continuing until the residual error is negligible.

The aspheric surface is represented by the expression

$$x = f(y, z) = \frac{cs^2}{1 + \sqrt{1 - c^2 s^2}} + A_2 s^2 + A_4 s^4 + \ldots \tag{9-143}$$

where $c = 1/R$, $s^2 = y^2 + z^2$. The first term is the equation for a spherical surface, and A_2, A_4, A_6, and so on, are the aspheric deformation constants of the surface.

The first step is to compute x_0, y_0, and z_0, the intersection coordinates of the ray with the sphere (of curvature c) which approximates the aspheric. This can be done through use of Eqs. (9-128) through (9-135). Then one calculates

$$\bar{x}_0 = f(y_0, z_0) \tag{9-144}$$

by stubstituting $s_0^2 = y_0^2 + z_0^2$ into Eq. (9-143). Then one computes

$$l_0 = (1 - c^2 s_0^2)^{1/2} \tag{9-145}$$

$$m_0 = -y_0 [c + l_0 (2A_2 + 4A_4 S_0^2 + \ldots)] \tag{9-146}$$

$$n_0 = -z_0 [c + l_0 (2A_2 + 4A_4 S_0^2 + \ldots)] \tag{9-147}$$

$$G_0 = l_0 (\bar{x}_0 - x_0)/(Xl_0 + Ym_0 + Zn_0) \tag{9-148}$$

$$x_1 = G_0 X + x_0 \tag{9-149}$$

$$y_1 = G_0 Y + y_0 \tag{9-150}$$

$$z_1 = G_0 Z + z_0 \tag{9-151}$$

The computations of Eqs. (9-144) through (9-151) are repeated, with subscripts increased by one each time until

$$\bar{x}_k = x_k \tag{9-152}$$

to within the accuracy required.

The refraction at the surface is calculated by

$$O^2 = l_k^2 + m_k^2 + n_k^2 \tag{9-153}$$

$$\bar{E}_1 = Xl_k + Ym_k + Zn_k \tag{9-154}$$

$$\bar{E}'_1 = [O^2(1 - \mu_1^2) + \mu_1^2 \bar{E}_1^2]^{1/2} \tag{9-155}$$

$$\bar{g}_1 = (\bar{E}'_1 - \mu_1 \bar{E}_1)/O^2 \tag{9-156}$$

$$X_1 = \mu_1 X + \bar{g}_1 l \tag{9-157}$$

$$Y_1 = \mu_1 Y + \bar{g}_1 m_k \tag{9-158}$$

$$Z_1 = \mu_1 Z + \bar{g}_1 n_x \tag{9-159}$$

Then the direction cosines X_1, Y_1, Z_1 and the intersection coordinates x_k, y_k, z_k define the refracted ray.

9.5.4. Graphical Ray Tracing. This technique applies Snell's law ($n \sin I = n' \sin I'$) at each surface of the system. It may be carried out by drawing a ray to its intersection with the surface, constructing the normal to the surface, measuring the angle I with a protractor, and then calculating the angle I'.

A purely graphic technique is shown in Fig. 9-23. The ray is drawn to the surface and the normal to the surface is erected at the point of intersection. Two circles are drawn about the point of intersection with radii proportional to the indices on either side of the surface. From the intersection of the ray with circle n at A, a line is drawn parallel to the normal to intersect circle n' at B; then the refracted ray is drawn through B and the ray-surface intersection.

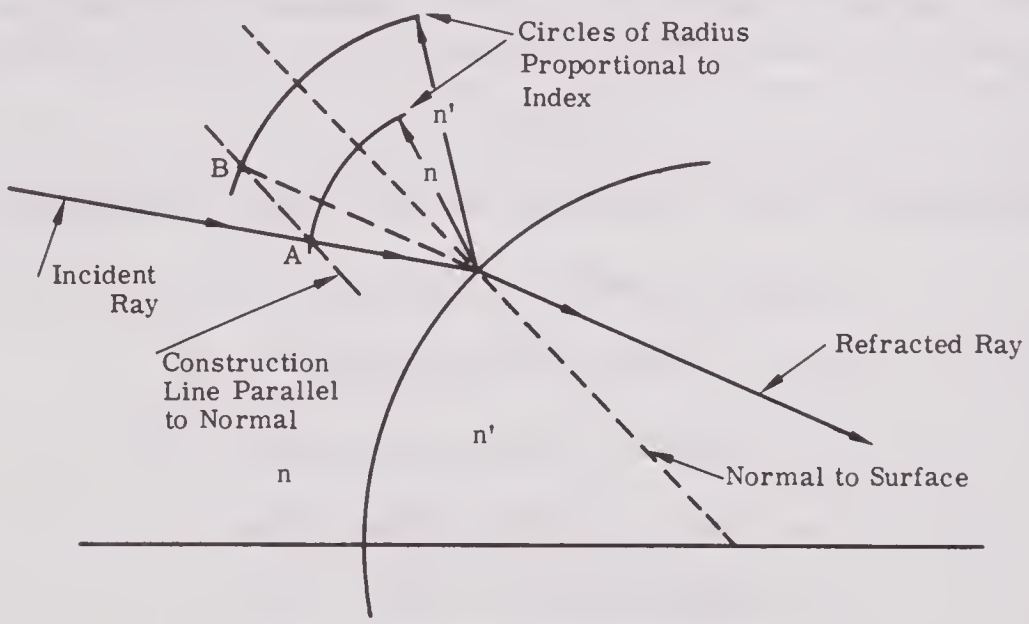


FIG. 9-23. Graphical ray tracing. Starting with the construction of circles (with radii proportional to the indices on either side of the surface) about the point of intersection of the ray and surface and the development of the refracted ray.

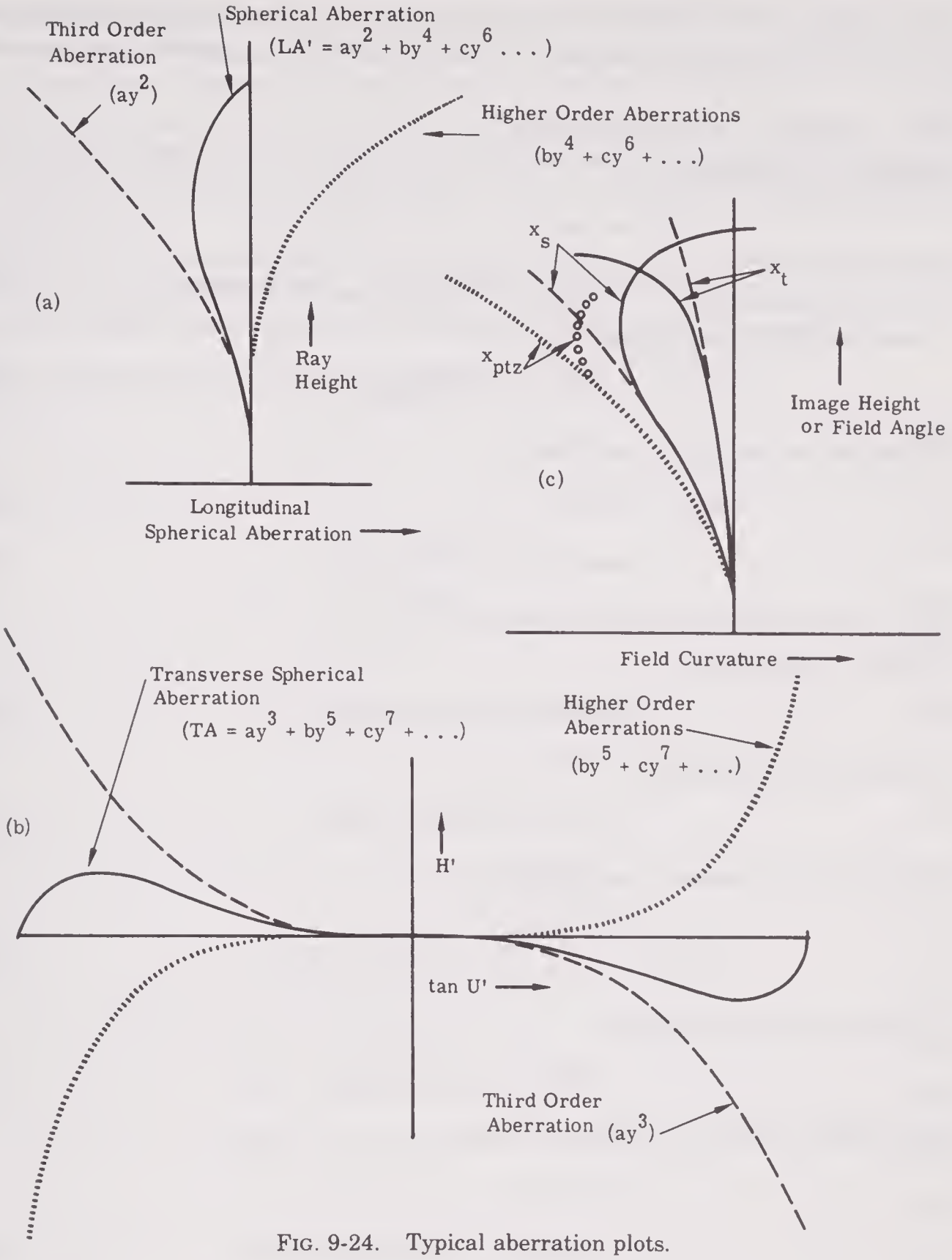


FIG. 9-24. Typical aberration plots.

9.6. Third-Order Aberrations

Figure 9-24 illustrates typical third-order aberrations.

9.6.1. Third- and Higher-Order Aberrations – Surface Contributions. The contribution of a given surface to the third-order aberration of the final image can be calculated by the equations in the following sections. The final third-order aberration is the summation of the contributions of all the surfaces. Two paraxial rays are traced through the system; one is the ray from the axial object point passing through the edge of the entrance pupil, a paraxial marginal ray, and the other is the ray through the center of the entrance pupil from an object point at the edge of the field, a paraxial principal ray. In the equations, quantities with the subscript *pr* denote the data of the paraxial principal ray; quantities without subscript refer to the data of the paraxial marginal ray.

The equations are evaluated for each surface, from the ray-trace data of that surface. The contributions are then summed to obtain the total third-order aberration of the system.

9.6.1.1. *Equations for Desk Calculator.*

Spherical contribution:

$$SC' = yni(i' - i)\,(i' - u)/2n'_k u'^2_k \tag{9-160}$$

$$SC' = yu'n'(u^2 - u'^2)/2n'_k u'^2_k \quad \text{(for a plane)} \tag{9-161}$$

Coma (sagittal) contribution:

$$CC'_s = SC'Ru'_k \tag{9-162}$$

Astigmatism (sagittal) contribution:

$$AC'_s = SC'R^2 \tag{9-163}$$

$$= ynui_p^2(n' - n)/n'2n'_k u'^2_k \tag{9-164}$$

(for use when object is near the center of curvature).

Petzval contribution:

$$PC' = (n - n')h_k^2 n'_k/2nn'r \tag{9-165}$$

Distortion contribution:

$$DC' = (AC' + PC')Ru'_k \tag{9-166}$$

Longitudinal chromatic contribution:

$$L\text{ch}C' = yni\left(\frac{\Delta n}{n} - \frac{\Delta n'}{n'}\right) \Big/ n'_k u'^2_k \tag{9-167}$$

Lateral color contribution:

$$T\text{ch}C' = (L\text{ch}C')Ru'_k \tag{9-168}$$

The symbols above are defined as follows:

$$R = i_p/i$$

$$= u_p/u \quad \text{(for a plane)} \tag{9-169}$$

SC' is the longitudinal third-order spherical aberration contribution.

CC' is the sagittal third-order coma contribution and is equal to one-third of the tangential coma contribution.

PC' is the contribution to the third-order Petzval curvature and $x_{\text{ptz}} = \Sigma PC'$.

$L\text{ch}C'$ is the axial longitudinal chromatic aberration and $\Sigma L\text{ch}c'$ is equal to $l'_v - l'_r$.

$T\text{ch}C'$ is the transverse off-axis chromatic contribution and $\Sigma t\text{ch}C' = h'_v - h'_r$.

AC' is the longitudinal sagittal astigmatism.

The third-order field curvatures may be found from

$$x_s = \Sigma PC' + \Sigma AC'$$

$$x_t = \Sigma PC' + 3\Sigma AC' \tag{9-170}$$

r is the radius of curvature, y is the ray height at the surface, n is the index of refraction, i is the angle of incidence, and u is the ray-slope angle; these are obtained from the ray-trace data. Rays may be traced by Eqs. (9-174) through (9-181). The data of the final image are primed and subscripted with k. Thus h'_k, u'_k, and n'_k refer to the final image height (*i.e.*, the height at which the paraxial principal ray strikes the final image plane), the final ray-slope angle at the image, and the index of the image space, respectively.

The dispersion of the medium is represented by $\Delta n = n_v - n_r$, where the subscripts refer to the short (v) and long (r) wavelengths of light. For visual work, F and C light (0.486 μ and 0.656 μ, respectively) are customarily used for these wavelengths.

9.6.1.2. *Equations for Electronic Computers or Desk Calculators.*

$$c = 1/\mathrm{r} \tag{9-171}$$

$$N = n/n' \tag{9-172}$$

$$u' = cy(1 - N) + Nu \tag{9-173}$$

$$y_{i+1} = y_i - t'_i u'_i \tag{9-174}$$

$$i = cy - u \tag{9-175}$$

$$h'_k = n_1(u_1 y_{p_1} - y_i u_{p_1})/\mathrm{n}'_k u'_k \tag{9-176}$$

$$I = n(uy_p - u_p y) = \text{invariant} \tag{9-177}$$

$$Z = (N - 1)c/n \tag{9-178}$$

$$B = ny(u' - i)(1 - N)/2I \tag{9-179}$$

$$B_p = ny_p(u'_p - i_p)(1 - N)/2I \tag{9-180}$$

$$TSC = Bi^2 h'_k \tag{9-181}$$

$$SC = TSC/u'_k \tag{9-182}$$

$$CC = Bii_p h'_k \tag{9-183}$$

$$TAC = Bi_p^2 h'_k \tag{9-184}$$

$$AC = TAC/u'_k \tag{9-185}$$

$$TPC = ZIh'_k/2 \tag{9-186}$$

$$PC = TPC/u'_k \tag{9-187}$$

$$DC = h'_k[B_p ii_p + (u'^2_p - u_p^2)/2] \tag{9-188}$$

$$TLC = yi(\Delta n - N\,\Delta n')/u'_k \tag{9-189}$$

$$L\mathrm{ch}C = TLC/u'_k \tag{9-190}$$

$$T\mathrm{ch}C = yi_p(\Delta n - N\,\Delta n')u'_k \tag{9-191}$$

The symbols above have the same meanings as in Sec. 9.6.1.1. The contributions TSC, TAC, TPC, and TLC are the transverse aberrations for spherical aberration, astigmatism, Petzval curvature, and longitudinal chromatic aberrations (which are customarily expressed as longitudinal aberrations). The transverse aberration is simply the longitudinal aberration times the final ray slope angle, u'_k.

Contributions from an aspheric surface: The aspheric surface is defined by

$$x = (1/2)c_e s^2 + [(1/8)c_e{}^3 + K]s^4 + \ldots \tag{9-192}$$

in which the terms in s^6 and higher may be neglected. c_e in this expression is not the same c as that used in Eq. (9-143) on aspheric ray tracing. For aspherics in the form of Eq. (9-143), an "equivalent" c_e and the equivalent fourth-order deformation constant K can be computed from

$$c_e = c + 2A_2 \tag{9-193}$$

$$K = A_4 - A_2(4A_2{}^2 + 6A_2c + 3c^2)/4 \tag{9-194}$$

The contributions are determined for the "equivalent" spherical surface by Eqs. (9-171) through (9-191). Then the additional contributions due to the "equivalent" fourth-order deformation constant K are computed by the following equations and added to those of the "equivalent" spherical surface to obtain the total third-order aberration contribution of the aspheric.

$$W = 4(n - n')K/I \tag{9-195}$$

$$TSC_a = Wy^4h'_k \tag{9-196}$$

$$CC_a = Wy^2y_p{}^2h'_k \tag{9-197}$$

$$TAC_a = Wy^3y_p{}^2h'_k \tag{9-198}$$

$$TPC_a = 0 \tag{9-199}$$

$$DC_a = Wyy_p{}^3h'_k \tag{9-200}$$

$$TLC_a = 0 \tag{9-201}$$

$$T\text{ch}A_a = 0 \tag{9-202}$$

9.6.2. Third-Order Aberration Contributions of Thin Lenses. If the thin-lens fiction is used (*i.e.*, assuming that the thickness of an element is zero), a useful set of aberration-contribution equations for a single element may be derived from the preceding equations. The procedure is to trace a paraxial marginal ray and a paraxial principal ray through the system using

$$u' = u + y\phi \tag{9-203}$$

$$y_{i+1} = y_i - d'_i u'_i \tag{9-204}$$

$$l'_k = y_k/u'_k \tag{9-205}$$

Then for each element

$$v = u/y \qquad (\text{or } v' = u'/y) \tag{9-206}$$

$$Q = y_p/y \tag{9-207}$$

Then the contributions may be determined from

$$SC^* = SC \tag{9-208}$$

$$CC^* = CC + SCQu'_k \tag{9-209}$$

$$AC^* = AC + CC\cdot 2Q/u'_k + SCQ^2 \tag{9-210}$$

$$PC^* = PC \tag{9-211}$$

$$DC^* = (PC + 3AC)Qu'_k + CC\cdot 3Q^2 + SCQ^3u'_k \tag{9-212}$$

$$L\mathrm{ch}\,C^* = L\mathrm{ch}C \tag{9-213}$$

$$T\mathrm{ch}\,C^* = L\mathrm{ch}CQu'_k \tag{9-214}$$

The starred terms are the contributions from an element which is not at the stop; that is, one for which $y_p \neq 0$. The unstarred terms are the contributions from the element when it is located at the stop (and $y_p = 0$) and are given by the following:

$$SC = -y^4(G_1C^3 - G_2C^2C_1 + G_3C^2v + G^4CC_1^2 - G_5CC_1v + G_6Cv^2)/u'^2_k$$
$$[\text{or } SC = -y^4(G_1C^3 + G_2C^2C_2 - G_3C^2v' + G_4CC_2^2 - G_5CC_2v' + G_6Cv'^2)/u'^2_k] \tag{9-215}$$

$$CC = -h'_k y^2(\tfrac{1}{4}G_5CC_1 - G_7Cv - G_8C^2)$$
$$[\text{or } CC = -h'_k y^2(\tfrac{1}{4}G_5CC_2 - G_7Cv' + G_8C^2)] \tag{9-216}$$

$$AC = -h'^2_k\phi/2 \tag{9-217}$$

$$PC = -h'^2_k\phi/2n = AC/n \tag{9-218}$$

$$DC = 0 \tag{9-219}$$

$$L\mathrm{ch}C = -y^2\phi/u'^2_k V \tag{9-220}$$

$$T\mathrm{ch}C = 0 \tag{9-221}$$

$$SSC = -y^2\phi P/u'^2_k V \tag{9-222}$$

The symbols not previously defined are G_1 through G_8, V, P, and SSC. The quantity V is the Abbé V number, the reciprocal relative dispersion of the material given by

$$V = (n-1)/\Delta n \tag{9-223}$$

The quantity P is the partial dispersion, given by

$$P = (n_m - n_r)/\Delta n \tag{9-224}$$

where n_r, n_m, and n_v are the indices at the long, middle, and short wavelengths, respectively, and $\Delta n = n_v - n_r$ as before. SSC is the contribution to the secondary spectrum of the system. Secondary spectrum is the residual longitudinal chromatic aberration when the foci for r and v light are equal; that is

$$SSC = l'_m - l'_r = l'_m - l'_v \tag{9-225}$$

The terms G_1 through G_8 are functions of the index of the element. (The thin-lens third-order equations are often called "G-sums.")

$$\begin{aligned} G_1 &= n^2(n-1)/2 & G_5 &= 2(n+1)(n-1)/n \\ G_2 &= (2n+1)(n-1)/2 & G_6 &= (3n+2)(n-1)/2n \\ G_3 &= (3n+1)(n-1)/2 & G_7 &= (2n+1)(n-1)/2n \\ G_4 &= (n+2)(n-1)/2n & G_8 &= n(n-1)/2 \end{aligned} \tag{9-226}$$

To apply the thin-lens aberration expressions, the contribution equations (9-208 through 9-214) are evaluated for each element of the system from the data of the thin-lens ray trace. The aberration at the image is then the sum of the contributions of all the elements.

For analytical work the thin-lens contribution equations are often set up so that the aberration contribution is expressed as a function of a parameter whose value is to be determined. For example, in a two-element system, the spherical and coma contributions might be expressed as functions of the curvature of the first surfaces of the elements (c_1 and c_3), giving

$$\begin{aligned} \Sigma SC &= SC_A + SC_B = ac_1^2 + bc_1 + d + ec_3^2 + fc_3 + g \\ \Sigma CC &= CC_A + CC_B = hc_1 + j + kc_3 + m \end{aligned} \tag{9-227}$$

if both elements are assumed to be in contact with the stop. If a zero value for the thin-lens third-order contribution were desired, the simultaneous solution of the two equations below would yield the necessary values of c_1 and c_3:

$$\begin{aligned} \Sigma SC &= 0 = ac_1^2 + bc_1 + ec_3^2 + fc_3 + (d + g) \\ \Sigma CC &= 0 = hc_1 + Kc_3 + (j + m) \end{aligned} \tag{9-228}$$

9.6.3. Stop Shift Theory. The aberration contribution of an element depends upon its position relative to the stop or pupil of the system.

9.6.3.1. *Spherical Aberration.* Spherical aberration is not affected by the position of the stop. It is a function of the effective size of the aperture, but the pupil may be placed anywhere in a system without changing the spherical aberration.

9.6.3.2. *Coma.* Coma is affected by a change in stop position if there is spherical aberration in the system. This is indicated (although not proved) by Eq. (9-209). The value of Q is a function of y_p (the height at which the principal ray strikes the element), which in turn is a function of the stop position. If the system has no spherical aberration, the stop position has no effect on coma.

9.6.3.3. *Field Curvature.* The stop position has no effect on the Petzval curvature of a system but does affect the astigmatism and thus the field curvature, if coma or spherical aberration are present, as indicated by Eq. (9-210). The astigmatism of a thin element at the stop is a function of the element power only and cannot be changed by bending.

9.6.3.4. *Distortion.* Distortion is affected by the stop position; a thin element in contact with the stop has no distortion contribution.

9.6.3.5. *Chromatic Aberration.* Longitudinal (axial) chromatic is not a function of stop position, whereas lateral color is a function of the stop position. A thin element at the stop has no lateral color contribution.

9.6.3.6. *Example 1.* Figure 9-25 shows a thin positive meniscus element. The portion of the element used by an oblique bundle of rays is moved further from the optical axis as the stop is moved farther from the lens. The rim ray curve of this lens is sketched in Fig. 9-26. The effect of shifting the stop along the axis is to select a different portion of the rim ray curve. With the stop in position A, the rim ray curve indicates outward flaring coma and inward curving tangential field. In position C the coma is inward flaring and the field is inward curving. Position B yields a system which is free of coma and which has a slightly backward-curving tangential field.

This diagram illustrates how spherical aberration and coma are related, and also illustrates a basic law of stop shift theory – in the presence of undercorrected spherical aberration the position of the stop which eliminates coma (the "natural" stop position) also produces the most backward-curving field possible.

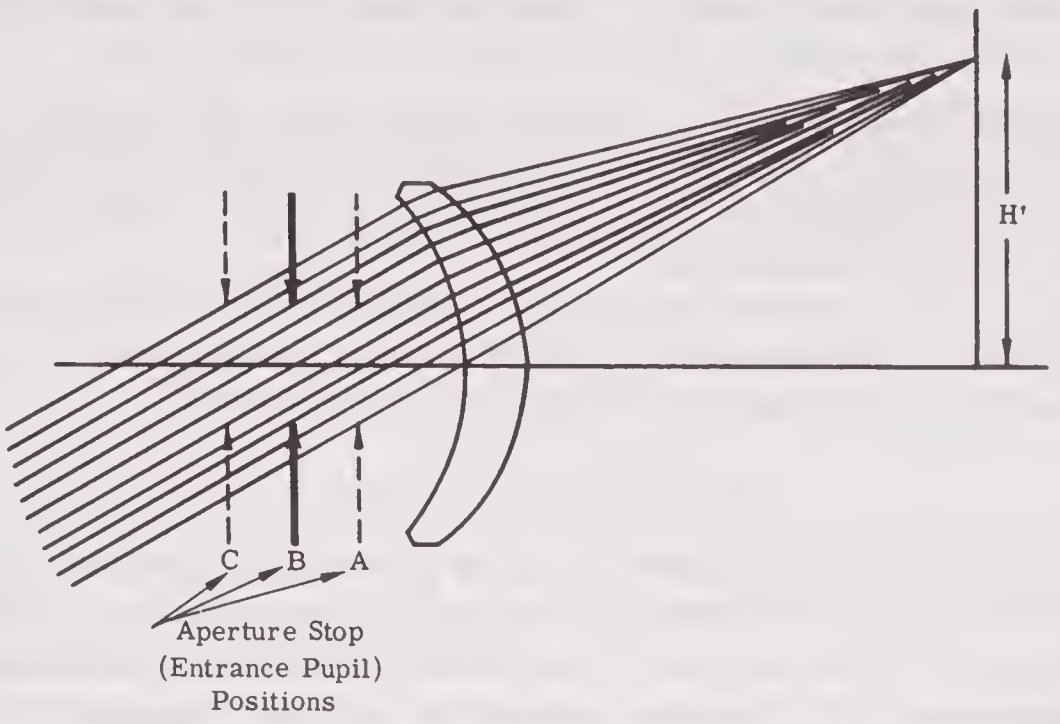


FIG. 9-25. Thin positive meniscus element.

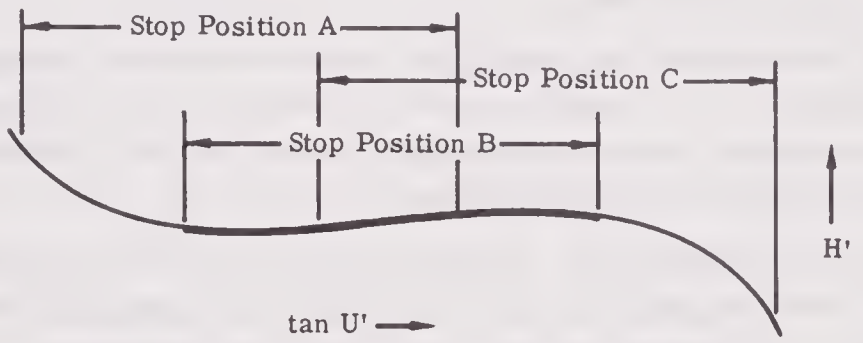


FIG. 9-26. Rim ray curve of lens in Fig. 9-25.

9.6.3.7. *Example 2.* Figure 9-27 shows a symmetrical erector system of the type used in terrestrial telescopes to invert the image formed by the objective. Each doublet is shaped to be free of spherical aberration; thus the size of the space between them, which determines the stop position, has no effect on the coma of the system. However, each doublet is designed to have a sizable amount of coma. Thus as the spacing between the doublets is changed, the astigmatism of the system is varied. The spacing is usually chosen so that the astigmatism is either zero or slightly positive so as to "artificially" flatten the field curvature.

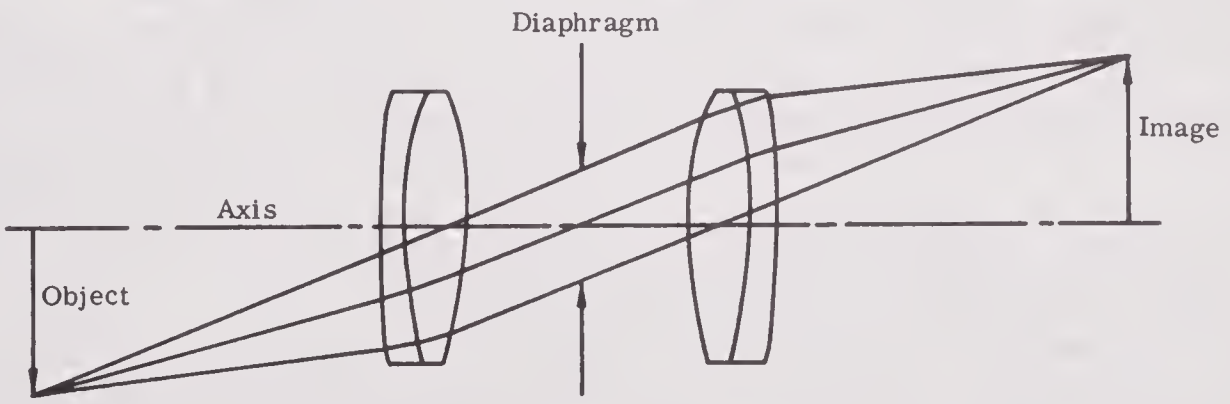


FIG. 9-27. Symmetrical erector system to invert image formed by objective.

9.6.3.8. *The Symmetrical Principle.* Example 2 is successful because the coma in each element is canceled by the coma in the other. In an optical system which is completely symmetrical about the stop, there is no coma, distortion, nor lateral color.

The provision of complete symmetry requires that the system work at unit magnification to be completely effective. However, even at infinite conjugates the coma, distortion, and lateral color are usually reduced to negligible values.

9.6.4. Afocal Systems. The expressions given above for third-order aberrations become indeterminate for afocal systems, in which u'_k is zero. Only a general outline of the necessary modifications to the equations is presented. (See Sec. 10.1.1 for a detailed discussion of afocal systems.)

In an afocal system, aberrations are best described in their angular form (See Sec. 9.4.1). Transverse aberrations (TA) and longitudinal aberrations (LA) can be converted to angular aberrations by use of the following equation:

$$\mathrm{AA} = \mathrm{TA}u'_k/y_1 = \mathrm{LA}u'^2_k/y_1 \tag{9-229}$$

The aberration contribution expressions of the preceding paragraphs can be modified in this manner to an angular aberration form; this will eliminate their interdeterminacy by canceling the u'_k which is found in one form or another in the denominators of each of these expressions. The following are longitudinal aberrations: $SC, AC, PC, L\mathrm{ch}C, SSC$. The following are transverse aberrations: $TSC, TAC, TPC, TLC, CC, DC, T\mathrm{ch}C$.

9.7. Optical Design Techniques

Optical design involves solving a number (say n) of second- (or higher) -order equations in m variables, where n represents the number of aberrations or characteristics which must be controlled and m represents the number of effective parameters that the designer has available for manipulation.

9.7.1. General Considerations. Primary requirements to be imposed on the optical system, aperture, focal length, and field coverage, and specialized requirements such as length of working distance must be determined. The resolution or definition necessary, and the spectral bandwidth must also be considered.

In infrared work, the choice of optical systems is usually between refracting and reflecting systems, depending on application, the materials acceptable to the application, and the necessity for chromatic correction.

9.7.2. Correction of Primary Aberrations. After the type of optical system has been selected or invented, the next major step in the design process is the correction of the primary aberrations, or at least the correction of as many of them as are necessary and feasible.

9.7.2.1. *First Steps.* The elements must be arranged to provide the desired optical characteristics, such as focal length, aperture, field, etc., for the system.

The usual method for correction of aberrations is bending of the elements. The longitudinal chromatic aberration, lateral color, Petzval curvature, and to a certain extent distortion are unaffected by bending. Chromatic aberration and Petzval curvature must be corrected in the original power and space layout. The thin-lens contribution equations (see Sec. 9.6.2) are useful at this stage, and it is ordinarily a relatively straightforward procedure to adjust the system so that the $\Sigma L\mathrm{ch}C$, $\Sigma T\mathrm{ch}C^*$, and ΣPC are equal to values which have been selected as desirable.

Then the spherical aberration, coma, astigmatism, and distortion must be corrected to their desired values. It is probably best at this stage to make a graph of the aberration contributions from each element as a function of the element shape. From a set

of such graphs a region (or regions) for the solution can be selected. These graphs can be made from data obtained by the use of the thin-lens contribution equations (see Sec. 9.6.2), the surface contribution equations (see Sec. 9.6.1) or, in certain cases, by direct ray tracing. The last two procedures are more appropriate for work with electronic computers.

When the region of the solution is selected, a method of differential correction is applied. The partial differentials of the aberrations against shape are computed and also the value of the aberrations for a trial prescription. The desired amount of change of each aberration (ΔA) is determined by analysis of the trial prescription; the necessary simultaneous equations of the form

$$\Delta A = \sum_{i=1}^{i=k} \frac{\partial A}{\partial C_i} \Delta C_i$$

are set up and solved. Because of the nonlinearity of the equations, the solution is seldom precise; however, the preselection of the solution neighborhood limits the size of ΔC so that the simultaneous solution is a good approximation and a series of solutions converges rapidly on the desired design shape.

9.7.2.2. *Limiting the Parameters.* If three aberrations A, B, and C are to be corrected by adjustment of three parameters x, y, and z, an initial trial of x, y, and z can be modified by changing one of the parameters, say z, so that one of aberrations, say C, is "corrected." Then parameter y is changed and a new value of z is determined to hold the correction of C. Parameter y is varied in this manner until aberrations B and C are simultaneously corrected. Then parameter x is changed and, with each change of x, y, and z are adjusted as above to hold aberration B and C as desired. Then x is varied in this manner until aberration A is brought to correction simultaneously with B and C. Graphs of C *vs.* z, B *vs.* y, and A *vs.* x are useful in such a procedure.

9.7.2.3. *Adding Thickness.* If the thin-lens expressions have been used in the preceding steps, it is necessary to add thickness to the elements. This is generally done by adjusting the secondary curvature of each thick element to hold the thick-element power equal to the thin-lens element power. The spacing between elements is then adjusted so that the separation of the principal points of the thick element is equal to the thin-lens spacings. This method serves to retain the overall system power and working distance at the same values as the thin-lens system.

9.7.2.4. *Trigonometric Corrections.* When the aberrations have been corrected using third-order aberration contributions (either thin lens or surface contributions), it is necessary to trace rays trigonometrically to determine the actual state of correction of the system. It will usually differ by a small amount from that predicted by the third-order expressions. However, a step or two of differential correction as outlined in Sec. 9.7.2.1 will usually bring the trigonometrical correction to the correct value.

It is possible to go directly to trigonometric correction from the thin-lens expressions (provided that the method of introducing thickness is rigorously consistent). Alternatively, an additional step of determining and correcting the surface contributions may be desirable after thicknesses are introduced.

9.7.3. Reduction of Residual Aberrations. After the primary aberrations have been brought to correction, the design is usually tested for residual aberrations. The primary aberrations are generally corrected for a single zone of the aperture or field and can be expected to depart from correction in all other zones, as discussed in Sec. 9.4.5.

If there were any parameters that were not used in the correction of the primary aberrations, these may be systematically varied and their effects on the residuals

noted and used. The possibility that more than one "neighborhood of solution" exists should not be overlooked; it is, in effect, an extra parameter.

An analysis of the source of the third-order contributions will often pinpoint one or two especially heavy contributory elements of the system. A reduction of a single large contribution will often reduce residual aberrations. This can be accomplished by introducing a correcting element near the offender (for example, convert a single positive element into a positive-negative doublet) or by splitting the offending element into two elements whose total power equals that of the original. This latter technique introduces two new variable parameters; the ratio of the powers of the two new elements, and the shape of the added element. Alternatively, a new shape for the offender may reduce its contribution to an acceptable level.

Where residuals are a problem, it is wise to reconsider the starting power and spacing layout. It is sometimes possible to revise the layout in such a way that the powers of the elements can be reduced. This is a rapid way of reducing residuals.

9.7.4. Automatic Design. The electronic computer has made possible the automatic implementation of certain of the preceding steps; for example correction of the third-order aberrations to a desired set of values. In this technique the computer calculates the partial differentials of the aberrations with respect to the available parameters and solves the resulting simultaneous equations to determine the required changes necessary. These approximate changes are then automatically applied to their respective parameters, and the process is repeated until the aberration contribution sums are within predetermined limits of the desired values. The system is then submitted to a trigonometric ray-trace check, and the process is repeated if necessary until the ray-trace aberrations are corrected.

Another school of automatic correction uses a "merit function," which is typically a weighted average of the absolute departures of the intersections of many rays from an ideal point image. Various techniques (least squares, steepest descent, and others) are used to improve the "merit function" automatically.

It is not possible in a handbook of this size to provide the automatic lens-design programs. From the references cited in the summary below, one can obtain some of the programs and get in touch with those responsible for their development or use.

Two distinct types of automatic lens correction are evident in the published literature. The first code gives the problem to the computer in explicit mathematical terms, thus making it possible for any engineer or scientist, with a modest knowledge of optics, to obtain lens designs [1-4]. The second type requires direction by a skilled specialist who makes qualitative judgments and compromises; hence the computer should be regarded as a tool presenting the designer with provisional solutions only [5,6]. Work is also being done in the design and manufacturing of aspheric optical elements [7,8]. Bell and Howell have been quite active in this branch of design and have developed a computer-programmed lens-grinding system [9].

Procedures and typical designs are presented in a thesis by G. Spencer [10], from which we quote directly:

> In 1954, Rosen and Eldert described a method designed to reduce the values of a large number of ray deviations – a number in excess of the number of available variables [11]. This led them to a least squares formulation. Hopkins, McCarthy, and Walters [12] and McCarthy [13], on the other hand, were interested in arriving at specific values for the first-order chromatic and third-order monochromatic aberrations – altogether seven in number and less than the number of available variables. This led them to a modified Newton-Raphson procedure, described by them in 1955. Feder phrased the problem in terms of the reduction of a single

"merit function" for the system with the implication that the smaller the value of the merit function the better the state of correction of the system. This led him to an extensive investigation of various gradient methods which he published in 1957 [14]. Modified gradient methods which maintain constant values for certain system characteristics while simultaneously reducing the value of a merit function have been described both by Feder [14] and by Meiron and Lobenstein [15].

The methods of Rosen and Eldert and of Hopkins and McCarthy may be termed linearization methods since they involve the approximation of non-linear functions by linear ones at each iteration. Practically all of the procedures thus far reported in the literature may be classified as either linearization methods or gradient methods. Procedures which defy these classifications have been described by Black [16] and by Meiron and Volinez [17]. These methods involve the successive adjustment of individual system parameters to values which minimize a figure of merit for the system. They may be called relaxation methods. Black also mentions the use of block and group operations on system parameters, which are standard relaxation techniques. Relaxation methods appear to be considerably less efficient than either gradient or linearization methods, however.

Probably the most successful automatic correction methods to appear thus far are the SLAMS method introduced by Wynne [18] and Nunn and Wynne [19] and the conjugate gradient method investigated by Feder [20].

Other references include the work of Holladay [21], of Gray [22], and of O'Brien [23].

9.8. Achromatism and Achromats

Many infrared systems make use of reflecting optics because of their freedom from chromatic aberration.

The condition for achromatism can be taken directly from the thin-lens aberration contribution equation (9-220) and is given by

$$\Sigma L\text{ch}C = 0 = -\frac{1}{U'_k{}^2}\,\Sigma y^2\phi/V \tag{9-230}$$

For a system of two refracting elements in close contact, the powers of the elements may be solved for directly, giving

$$\phi_A = \frac{V_A}{(V_A - V_B)}\,(\phi_{AB} - RV_B) \tag{9-231}$$

$$\phi_B = \frac{V_B}{(V_B - V_A)}\,(\phi_{AB} - RV_A) \tag{9-232}$$

$$R = \frac{L\text{ch}A}{l^2} - \frac{L\text{ch}A'}{l'^2} = \left(\frac{l_v - l_r}{l_v l_r}\right) - \left(\frac{l'_v - l'_r}{l'_v l'_r}\right) \tag{9-233}$$

where ϕ_A, ϕ_B, ϕ_{AB} are the powers of the elements A and B and the doublet AB, respectively, V_A and V_B are the Abbé V numbers $(n_m - 1/n_v - n_r)$ of the elements A and B, and R is a residual chromatic aberration term as defined in Eq. (9-233). In the Eq. (9-233), $L\text{ch}A$ is the chromatic aberration of the object a distance l from the doublet and $L\text{ch}A'$ is the chromatic desired in the image located a distance l' from the doublet: l_v and l_r are the object distances in short and long wavelengths, respectively, and l'_v and l'_r are the image distances. If a real object is to be imaged without any chromatic aberration (as is usually the case) the equations reduce to

$$\phi_A = V_A\phi_{AB}/(V_A - V_B) \tag{9-234}$$

$$\phi_B = V_B\phi_{AB}/(V_B - V_A) \tag{9-235}$$

In the infrared, it is necessary to check the secondary spectrum carefully. For example, ordinary optical glasses are often used in the lead sulfide region and one might calculate an "achromat" from Eqs. (9-234) and (9-235) using V-values based on wavelengths of 1.0 μ, 1.8 μ, and 2.5 μ which would yield a doublet that brought radiation of 1.0 μ and 2.5 μ to a common focus. However, on ray tracing the intermediate wavelengths one would find a very large secondary spectrum, because for every pair of optical glasses there is a wavelength between 1.0 μ and 1.5 μ at which their reciprocal relative dispersion (V values) is identical. At this wavelength the doublet is no better corrected than a single element. To achieve chromatic correction over this particular spectral region it is necessary to use fluorite (CaF_2) in combination with a dense barium crown or a light flint glass.

The thin-lens expression for secondary spectrum (Eq. 9-222) indicates a technique for handling this problem. For the secondary spectrum to be zero, the conditions that $\Sigma L\text{ch}C = 0$ and $\Sigma SSC = 0$ are necessary. In a thin doublet

$$\Sigma SSC = -f^2\left(\frac{P_A\phi_A}{V_A} + \frac{P_B\phi_B}{V_B}\right) \tag{9-236}$$

and if values for ϕ_A and ϕ_B from Eqs. (9-234) and (9-235) are substituted into this expression, the following equation can be derived:

$$\Sigma SSC = f(P_B - P_A)/(V_A - V_B) \tag{9-237}$$

in which P_A must equal P_B to achieve a zero secondary spectrum. A plot of P against V for the available materials (P and V calculated for the spectral band of interest) is useful in selecting the pair of materials with the smallest value of $(P_B - P_A)/(V_A - V_B)$ and hence the smallest secondary spectrum.

9.9. Resolution, Definition, and Image Spot Size

The resolving power of a system is the smallest angular separation of two equally bright point sources at which the system can detect that there are two sources rather than one. The resolution does not fully describe the performance of an optical system, and other criteria such as energy distribution and frequency response have come into use. The energy distribution is often represented by a plot of the percentage of the total image energy falling within a circle of a given diameter against the diameter of the circle. Frequency response is the percentage modulation of intensity in the image of an object of a given spatial frequency and is usually presented as a plot of percentage modulation or response against spatial frequency.

9.9.1. The Airy Disc. The wave nature of light limits the characteristics of the image formed, even by an optical system without aberration, to a disc of illumination surrounded by rings of illumination which are progressively fainter for larger rings. This pattern is shown in Fig. 9-28. Figure 9-29 indicates the illumination levels in the pattern. (The illumination levels in the rings are highly exaggerated for clarity.)

For a circular aperture,* the illumination distribution in this pattern is given by

$$E = K\rho^4\left[1 - \frac{1}{2}\left(\frac{m}{2}\right)^2 + \frac{1}{3}\left(\frac{m^2}{2^2 2!}\right)^2 - \frac{1}{4}\left(\frac{m^3}{2^3 3!}\right)^2 + \frac{1}{5}\left(\frac{m^4}{2^4 4!}\right)^2 \cdots\right]^2 \tag{9-238}$$

*See Fig. 9-30 for meaning of symbols.

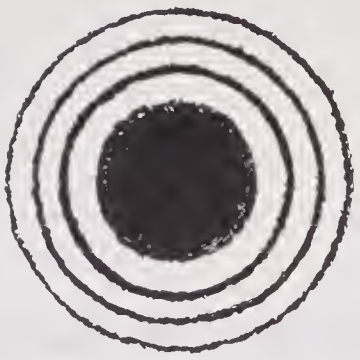


FIG. 9-28. Appearance of diffraction pattern.

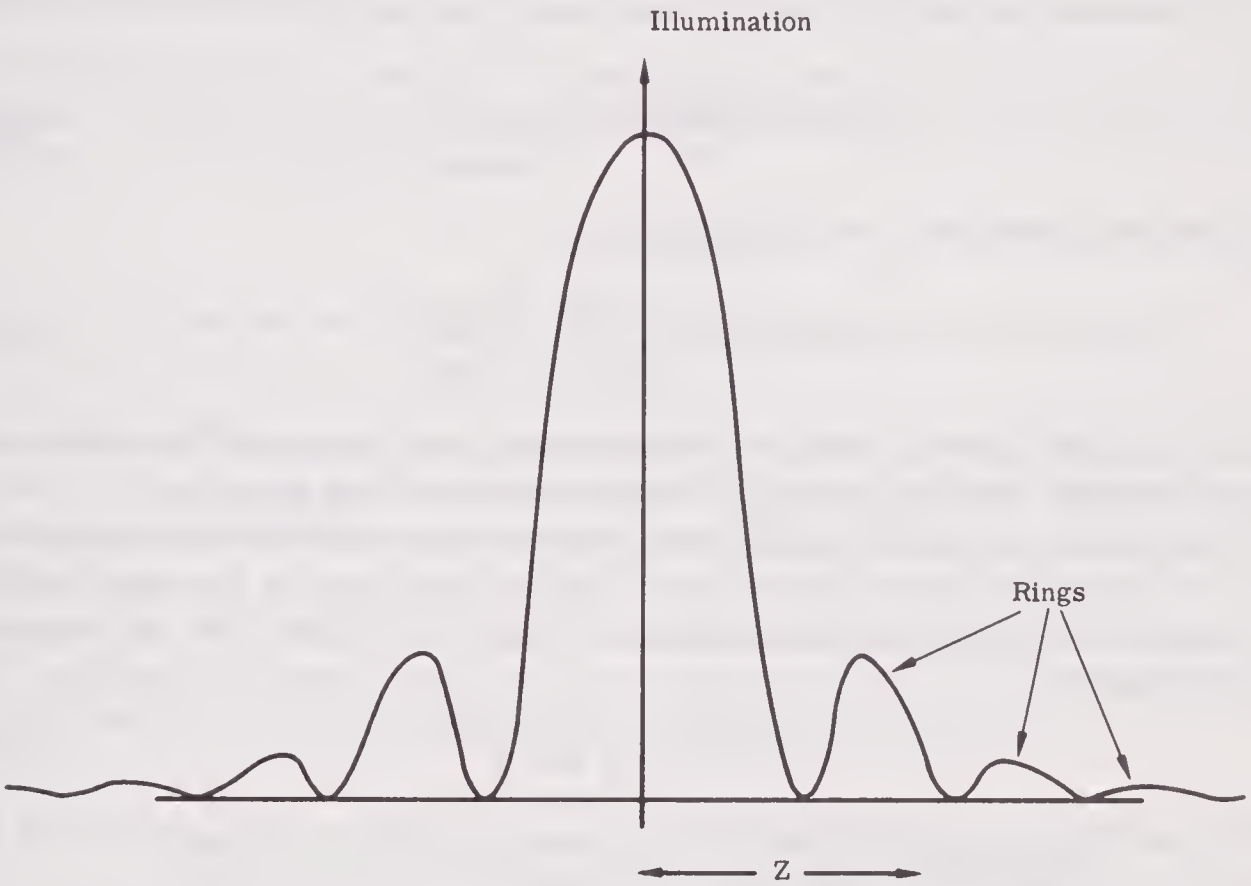


FIG. 9-29. Relative illumination in Airy disc with the size of the central maxima reduced in scale.

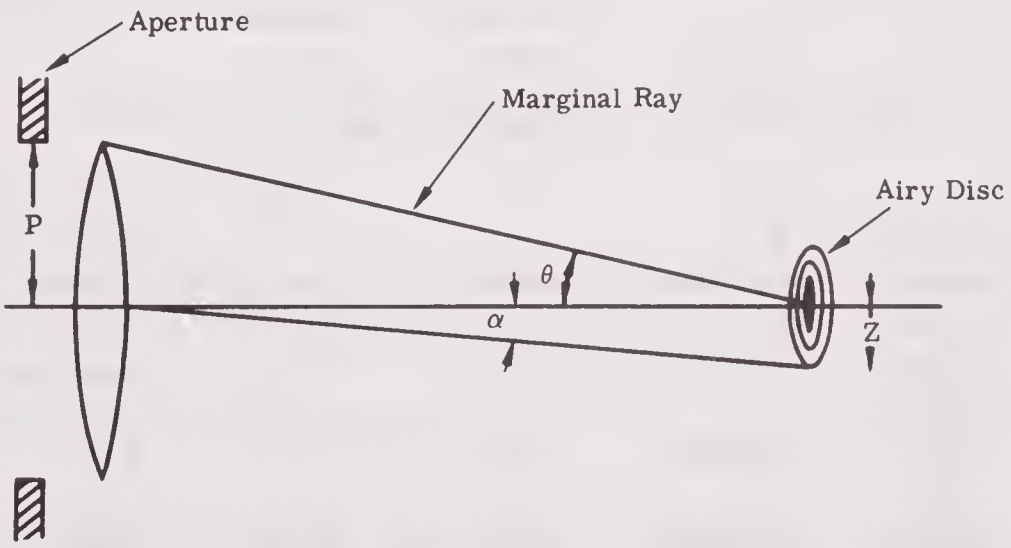


FIG. 9-30. Meaning of symbols in Eqs. (9-238) through (9-242) and Table 9-2.

in which

$$m = \frac{2\pi\rho}{\lambda} \sin \alpha \tag{9-239}$$

where E = the illumination at point Z

ρ = the semidiameter of the aperture of the system

λ = the wavelength of the energy

α = the angle subtended by the distance from the point Z to the axis, from the aperture

For a slit aperture the pattern becomes a series of lines and the illumination is given by

$$E = K \frac{\sin^2 m}{m^2} (1 + \cos \alpha) \tag{9-240}$$

For a rectangular aperture,* the illumination is

$$E = K\rho_1^2\rho_2^2 \frac{\sin^2 m_1}{m_1^2} \cdot \frac{\sin^2 m_2}{m_2^2} \tag{9-241}$$

Table 9-2 gives the location, relative illumination, and integrated illumination for the central disc and the first four rings for the circular and slit apertures.

The Rayleigh criterion assumes that two points are just resolved if the central maximum of one image pattern is directly over the first dark ring of the other, and *vice versa.* Figure 9-31 illustrates this condition. Thus (from Table 9-3) the separation for this condition is

$$Z = 0.61\lambda/n \sin \theta \tag{9-242}$$

where Z is the linear separation between the images, λ is the wavelength, n is the final index, and θ is the half-angle subtended by the aperture from the image.

TABLE 9-2. LOCATION, RELATIVE ILLUMINATION, AND INTEGRATED ILLUMINATION FOR CIRCULAR AND SLIT APERTURES

Ring (or Band)	*Circular Aperture*				*Slit Aperture*	
	Z ($\lambda/n \sin \theta$)	*Peak Illumination*	*Total Ring Energy Relative to Central Disc*	*Proportion of Total Energy in Each Ring*	*Z* ($\lambda/n \sin \theta$)	*Peak Illumination*
Central max.	0	1.0	1.0	0.839	0	1.0
1*st* dark	0.61	0	–	–	0.5	0
2*nd* bright	0.819	0.01745	0.084	0.071	0.715	0.0469
2*nd* dark	1.116	0	–	–	1.0	0
3*rd* bright	1.333	0.00415	0.033	0.028	1.230	0.0168
3*rd* dark	1.619	0	–	–	1.5	0
4*th* bright	1.847	0.00165	0.018	0.015	1.736	0.0083
4*th* dark	2.120	0	–	–	2	0
5*th* bright	2.361	0.00078	0.011	0.01	2.235	0.0050
5*th* dark	2.622	0	–	–	2.5	0

*See Fig. 9-30 for meaning of symbols.

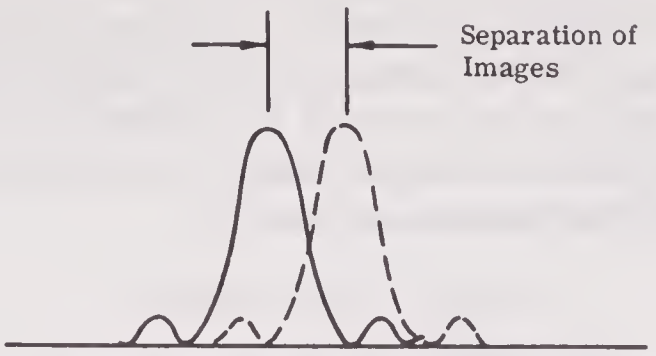


FIG. 9-31. Rayleigh criterion for resolution; peak intensity from one Airy disc dislocated at first dark ring of other Airy disc.

If the optical system is reasonably aplanatic, ρ/l' can be substituted for $n \sin \theta$ (if the image is in air) and the following expression derived for the limiting angular resolution (Z/l') of an optical system

$$\beta = 0.61\lambda/\rho \tag{9-243}$$

where β = the angular separation of the two objects which are just resolved

ρ = the semiaperture of the optical system

9.9.2. The Effects of Aberrations on the Airy Disc. Rayleigh established a criterion for the amount of aberration which could be tolerated without "sensibly" degrading the image of an otherwise perfect system. The Rayleigh limit can be expressed as follows.

An image can be "sensibly" perfect if there exists not more than one quarter-wavelength difference in optical path over the wavefront with reference to a spherical wavefront about the selected image point (see Tables 9-3 and 9-4).

TABLE 9-3. DISTRIBUTION OF ENERGY IN CENTRAL DISC

Amount of Aberration	*Energy in Central Disc* (%)	*Energy in Rings* (%)
Perfect system	84	16
1/4 Rayleigh limit	83	17
1/2 Rayleigh limit	80	20
1 Rayleigh limit	68	32

TABLE 9-4. RAYLEIGH LIMIT IN GEOMETRIC TERMS

One Rayleigh limit of:	
Out of focus =	$\lambda/2n' \sin^2 U'_m$
Third-order longitudinal spherical aberration =	$4\lambda/n' \sin^2 U'_m$
Residual zonal longitudinal spherical aberration =	$6\lambda/n' \sin^2 U'_m$
Sagittal coma =	$\lambda/2n' \sin U'_m$
Longitudinal chromatic aberration ~	$\lambda/n' \sin^2 U'_m$

9.9.3. Geometrical Limits on Resolution. When aberrations are small, their effects on image size should be evaluated in terms of the diffraction pattern. However, when the aberrations are quite large, the blur spot predicted by ray tracing may be used to determine the image characteristics.

9.9.3.1. *Third-Order Spherical Aberration.* Figure 9-32 shows the pattern of ray intersections at the image of a lens system with pure third-order spherical aberration, that is, aberration which can be described by the equation $LA = \alpha y^2$. There is a well-defined minimum diameter blur spot, which occurs at $0.75\ LA_m$ from the paraxial focus. At this point the blur diameter is given by

$$B_{min} = 0.5 LA_m \tan U'_m \tag{9-244}$$

and is one-fourth the size of the blur at the paraxial focus.

When the aberration is small, the position of best focus (chosen on the basis of wave-front aberration or minimum optical path difference, O.P.D.) is halfway between the mariginal focus and the paraxial focus.

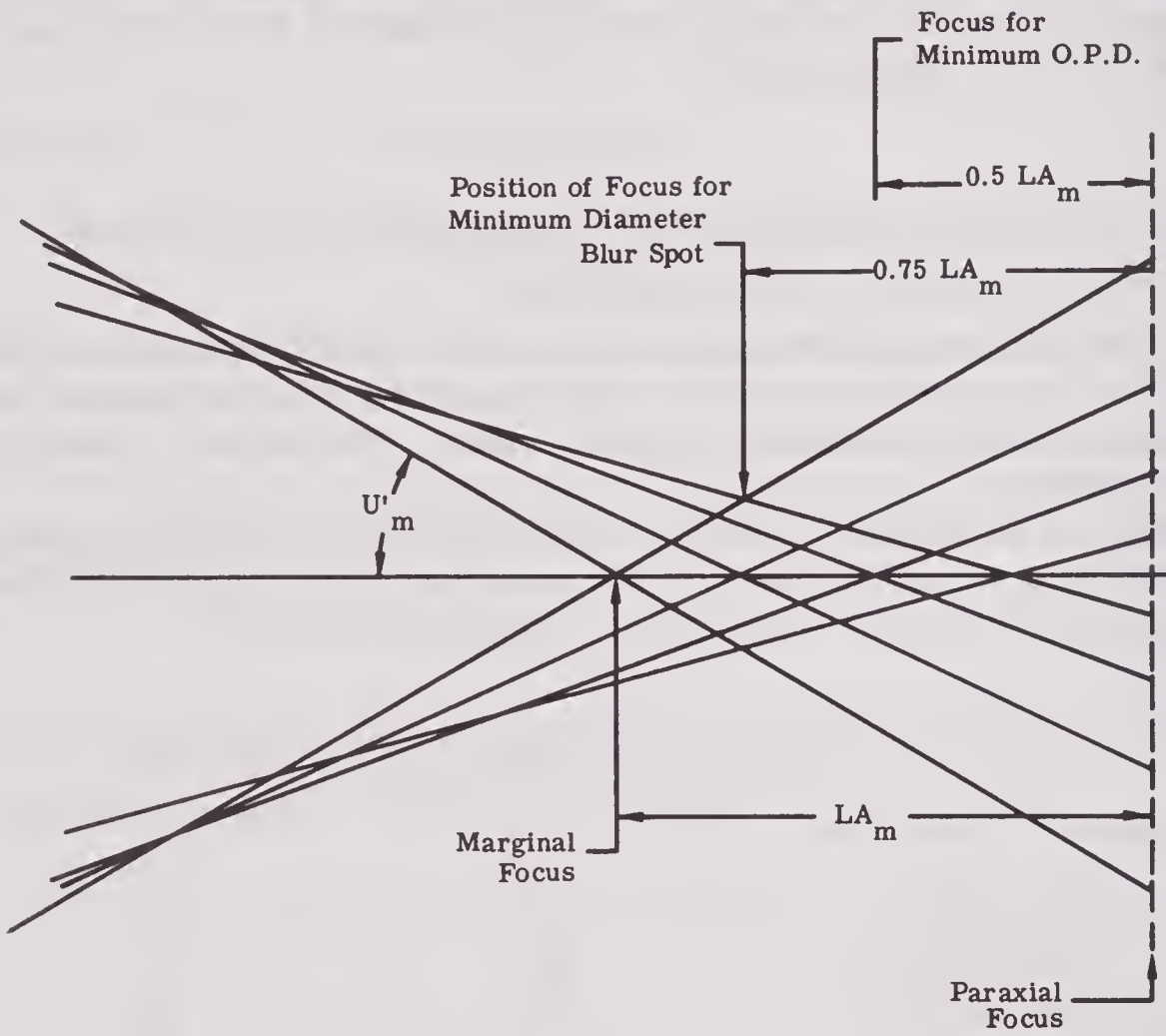


FIG. 9-32. Pattern of ray intersections at image lens system with pure third-order spherical aberration.

9.9.3.2. *Third- and Fifth-Order Spherical Aberration.* In the presence of third- and fifth-order spherical aberration (*i.e.,* when $LA = ay^2 + by^4$), there are two possibilities for optimization. On the basis of O.P.D. the best image occurs when the lens is designed so that $LA_m = 0$, but the minimum geometrical blur circle occurs when $LA_z = 1.5\, LA_m$. Both cases are illustrated in Fig. 9-33.

Marginal spherical equal to zero. The minimum O.P.D. focus is $0.75\, LA_z$ from the paraxial focus but the minimum diameter blur occurs at $0.422\, LA_z$ from the paraxial focus. At this latter point the blur diameter is given by:

$$B_{min} \approx 0.84 LA_z \tan U'_m \text{ (for small angles)}$$
$$B_{min} \approx 0.42 LA_z (\tan U'_m + \sin U'_m) \text{ for larger angles} \tag{9-245}$$

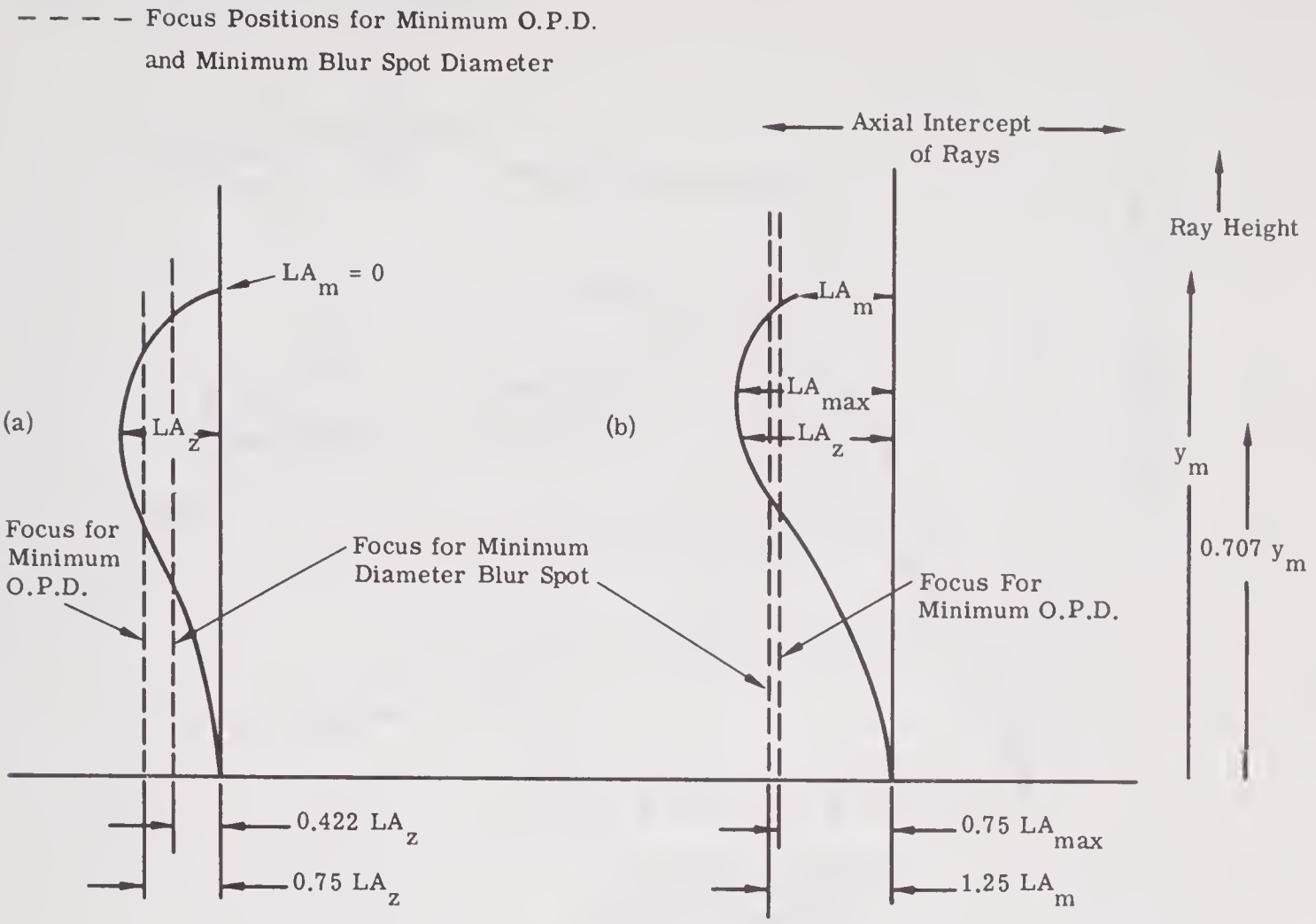


FIG. 9-33. Geometrical image formation in presence of third- and fifth-order spherical aberration, (*a*) where $LA_m = 0$ and (*b*) where $LA_z = 1.5\ LA_m$ (state of correction actually produces smaller geometrical blur spot although O.P.D. is less in case where $LA_m = 0$).

Marginal spherical equal to 0.67 zonal spherical. The minimum O.P.D. focus is $0.75\,LA_{max}$ from the paraxial focus but the minimum diameter of the blur spot is $1.25\,LA_m$ from the paraxial focus and is given by

$$B_{min} \approx 0.5\,LA_m \tan U'_m \tag{9-246}$$

9.9.3.3. *Chromatic Aberration.* The treatment of blur-spot size in the presence of longitudinal chromatic aberration is complicated because the spectral response of the detector is a factor. Figure 9-34 shows the image energy distribution under these circumstances. The insert figures show the spectral response plots. The main graphs plot the percentage of the total energy within a given diameter area, where the diameter of the area is given as a fraction of the total size of the chromatic blur spot. The total size of the chromatic blur spot can be determined from

$$B_{min} = LchA \cdot \tan U'_m = (l'_v - l'_r) \tan U'_m \tag{9-247}$$

The energy density in the chromatic blur spot is not uniform; 75 to 95% of the "effective" energy is contained within the central half diameter (or quarter of the area) of the spot.

It is often of interest to determine the image characteristics at the image of a sharp edge. The intensity at the image of an edge in the presence of chromatic for the two response characteristics is shown in Fig. 9-35.

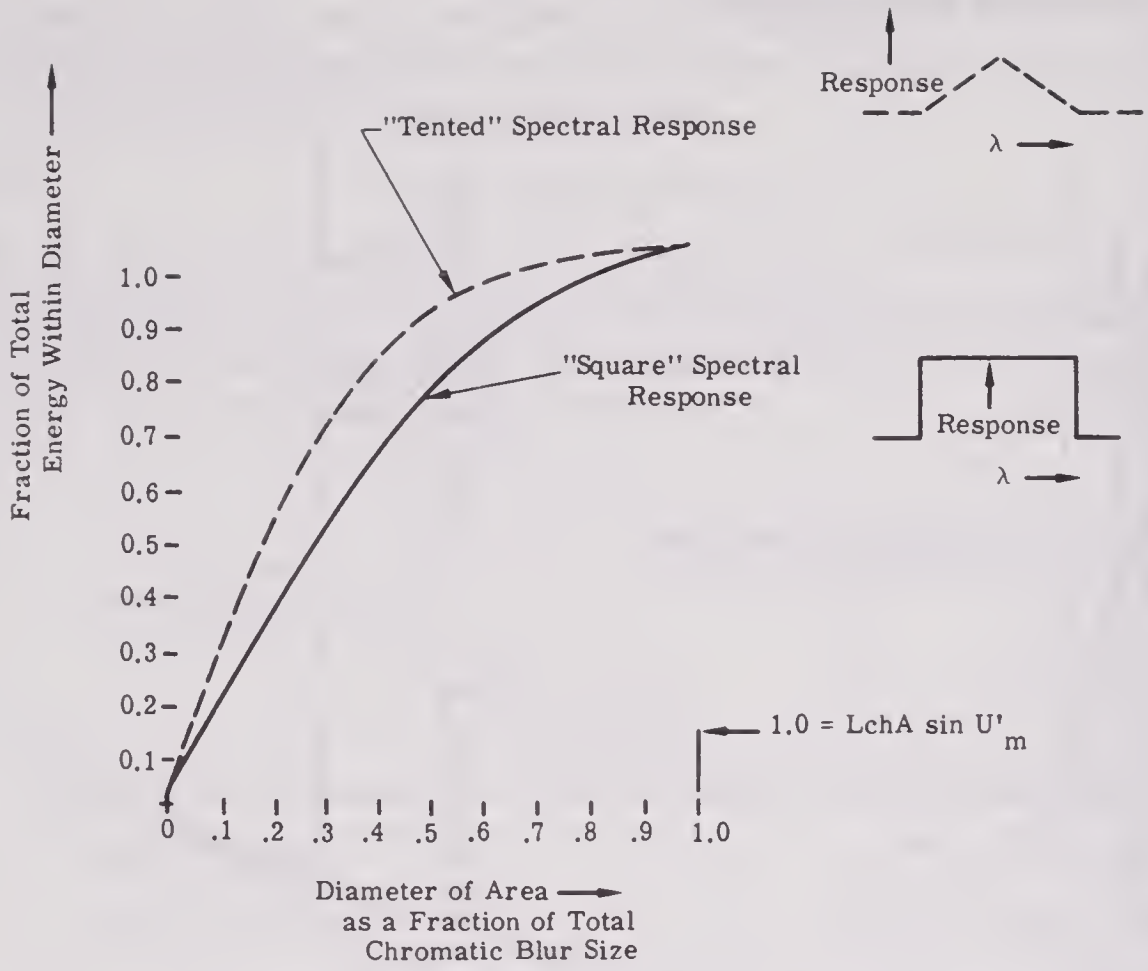


FIG. 9-34. Energy distribution in presence of longitudinal chromatic aberration with plane of reference midway between extreme focii. Graph indicates percentage of total energy in image that falls within a circle of given distance.

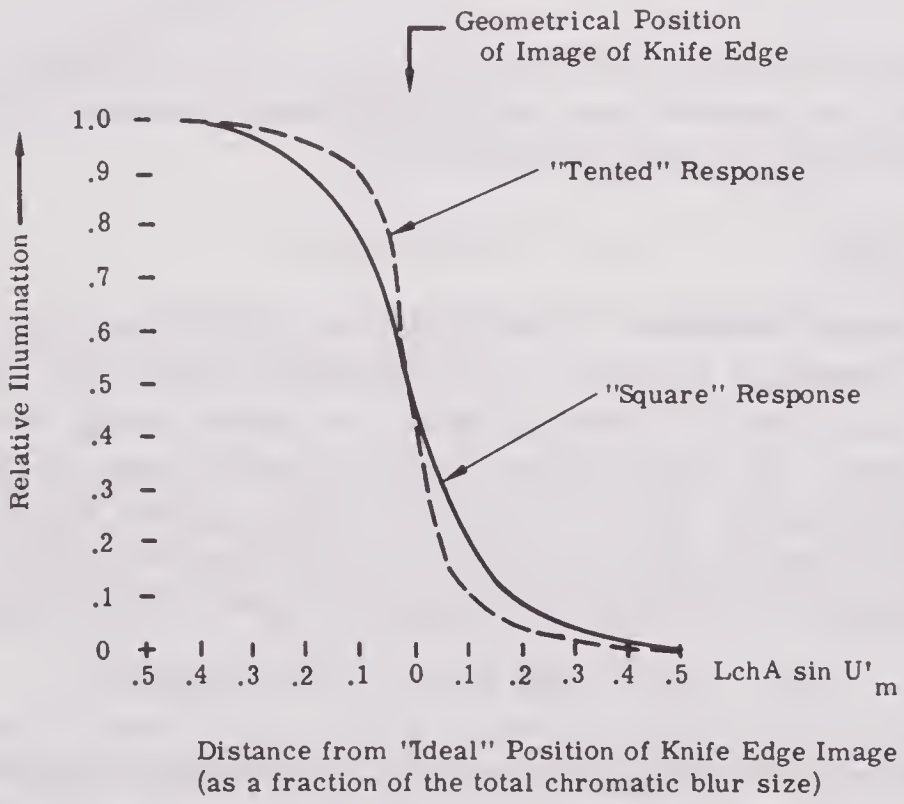


FIG. 9-35. Intensity gradient at image of knife edge in presence of longitudinal chromatic with plane of focus midway between extreme focii.

9.10. Summary of Equations

$$n' \sin I' = n \sin I$$

$$\frac{1}{s'} = \frac{1}{s} + \frac{1}{f}$$

$$f^2 = -xx'$$

$$m = -\frac{f}{x} = \frac{-x'}{f}$$

$$I = hnu = h'n'u' = ynu_{pr}$$

$$\text{power} = \phi = \frac{1}{f} = (n-1)\left[\frac{1}{r_1} - \frac{1}{r_2} + \frac{(n-1)t}{nr_1r_2}\right]$$

$$= (n-1)\left[c_1 - c_2 + \frac{(n-1)}{n} tc_1c_2\right]$$

$$u'_i = u_i + y_i\phi_i$$

$$y_{i+1} = y_i - d'_i u_i$$

$$f_{AB} = \frac{f_A f_B}{f_A + f_B - d}$$

$$f_A = \frac{f_{AB} d}{f_{AB} - s'}$$

$$f_B = \frac{-s'd}{f_{AB} - s' - d}$$

$$d = \frac{f_B s'}{f_B - s'} = f_A + f_B - \frac{f_A f_B}{f_{AB}} = \frac{(v + f_A) \pm \sqrt{(v - f_A)^2 - 4f_B(v - f_A)}}{2}$$

$$\Sigma\phi = \min$$

$$d = v - f_{AB} \pm \sqrt{f_{AB}(f_{AB} - v)}$$

$$\Sigma\phi = \min$$

$$s' = f_{AB} \pm \sqrt{f_{AB}(f_{AB} - v)}$$

$$(u'_i n'_i) = u_{i-1} n_i + y_i(n'_i - n_i)/r_i$$

$$y_{i+1} = y_i - t'_i(u'_i n'_i)/n'_i$$

$$\text{numerical aperture} = \text{N.A.} = n \sin u$$

$$LA = \frac{TA}{\tan U'_m} = \frac{AA \cdot f}{\tan U'_m}$$

$$LA' = L' - l'$$

$$\text{coma}_t = H'_{AB} - H'_{pr}$$

$$\text{O.S.C.} = \frac{\text{coma}_s}{h} = \frac{\sin U_1}{u_1} \cdot \frac{u'_k}{\sin U'_k} \cdot \frac{l' - l'_{pr}}{(L' - l'_{pr})} - 1$$

$$X_s = L'_{CG} - l'$$

$$X_T = L'_{AE} - l'$$

$$\text{Distortion} = H'_{pr} - h'$$

$$L\text{ch}A = l'_v - l'_r$$

$$T\text{ch}A = H'_v - H'_r$$

$$\text{C.D.M.} = (H'_v - H'_r)/h'$$

$$LA' = L' - l'$$

$$TA' = LA' \tan U'$$

$$\text{O.S.C.} = \frac{\sin U_1}{u_1} \cdot \frac{U'}{\sin U'} \cdot \frac{(l' - l'_{pr})}{(L' - l'_{pr})} - 1$$

$$x_{\text{ptz}} = 1.5x_s - 0.5x_t$$

$$\text{Dist} = H'_{pr} - h'$$

$$L\text{ch}A = l'_v - l'_r$$

$$T\text{ch}A = H'_v - H'_r$$

Desk Calculations

Paraxial Ray Trace: Sec. 9.5.3.1

Meridional finite radii: Eq. (9-73)

Meridional planes: Eq. (9-83)

Computer Calculations

Meridional: Eq. (9-101)

Skew ray, spherical surface: Eq. (9-128)

Skew ray, aspheric: Eq. (9-144)

Third-order surface aberrations: Eq. (9-160)

Third-order thin lens aberrations: Eq. (9-203)

$$\Sigma L\text{ch}C = 0 = -\frac{1}{U'_k{}^2} \Sigma y^2 \phi/V$$

$$\phi_A = \frac{V_A}{(V_A - V_B)} (\phi_{AB} - RV_B)$$

$$\phi_B = \frac{V_B}{(V_B - V_A)} (\phi_{AB} - RV_A)$$

$$R = \frac{L\text{ch}A}{l^2} - \frac{L\text{ch}A'}{l'^2} = \left(\frac{l_v - l_r}{l_v l_r}\right) - \left(\frac{l'_v - l'_r}{l'_v l'_r}\right)$$

$$\phi_A = V_A \phi_{AB}/(V_A - V_B)$$

$$\Sigma SSC = f(P_B - P_A)/(V_A - V_B)$$

$$E = K\rho_1{}^2\rho_2{}^2 \frac{\sin^2 m_1}{m_1{}^2} \cdot \frac{\sin^2 m_2}{m_2{}^2}$$

$$Z = 0.61\lambda/n \sin \theta$$

$$\beta = 0.61\lambda/\rho$$

References

1. C. A. Lehman, *LASL Design Program for the IBM 7090,* Los Alamos Scientific Laboratory Report LA-2837 (1963).
2. B. Brixner, *Lens Designing Technique with the 1962 LASL Code for the IBM 7090,* Los Alamos Scientific Laboratory Report LA-2877 (1963).
3. D. P. Feder, *J. Opt. Soc. Am.* **52,** 177 (1962).
4. Unpublished program of Illinois Institute of Technology Research Institute, Chicago, Ill.
5. R. E. Hopkins, *Research on Fundamentals of Geometrical Optics* (ORDEALS FOR IBM 7070), The Institute of Optics, University of Rochester, Contract AF-49(638)-668, Project 976, Task 37650 (1962).
6. R. E. Hopkins and G. Spencer, *J. Opt. Soc. Am.* **52,** 172 (1962).
7. G. W. Shavlik, *Modifications and Additions to the NOTS General Optical Ray Tracing Computer Program,* U.S. Naval Ordnance Test Station, China Lake, Calif., Report 7966 (1962).
8. J. R. Snyder, Photophysics Branch, Instrument Development Division, U.S. Naval Ordnance Test Station, China Lake, Calif.
9. D. Stevens and J. Thiel, AF 33(600)-37199 Bell & Howell Company, Chicago, Ill.
10. G. Spencer, Thesis, College of Engineering and Applied Science, University of Rochester, Rochester, N.Y.
11. S. Rosen and C. Eldert, *J. Opt. Soc. Am.* **44,** 250 (1954).
12. R. E. Hopkins, C. A. McCarthy, and R. Walters, *J. Opt. Soc. Am.* **45**, 363 (1955).
13. C. A. McCarthy, *J. Opt. Soc. Am.* **45,** 1087 (1955).
14. D. P. Feder, *J. Opt. Soc. Am.* **47,** 902 (1957).
15. J. Meiron and H. M. Lobenstein, *J. Opt. Soc. Am.* **47,** 1104 (1957).
16. G. Black, *Proc. Phys. Soc.* (London) **B68,** 729 (1955).
17. J. Meiron and G. Volinez, *J. Opt. Soc. Am.* **50,** 207 (1960).
18. C. G. Wynne, *Proc. Phys. Soc.* (London) **73**, 777 (1959).
19. M. Nunn and C. G. Wynne, *Proc. Phys. Soc.* (London) **74**, 316 (1959).
20. D. P. Feder, *J. Opt. Soc. Am.* **52,** 177 (1962).
21. J. C. Holladay, *Computer Design of Optical-Lens Systems (IBM 704) in Computer Applications – 1960,* Benjamin Mitteman and Andrew Ungar, (eds.) Macmillan Company, New York (1961).
22. D. S. Gray, *J. Opt. Soc. Am.* **53**, 672 (1963).
23. Private communication from C. O'Brien of Illinois Institute of Technology Research Institute.

Bibliography

The literature does not abound with books written by lens designers. They seem to design lenses. Some useful books are:

Born, M., and Wolf, E., *Principles of Optics,* Pergamon, New York (1959).

Buchdahl, H. A., *Optical Aberration Coefficients,* Oxford University Press, London (1954).

Conrady, A. E., *Applied Optics and Optical Design, Vols. I and II,* Dover, New York (1957 and 1959). This is the book on which most lens designers first cut their teeth, even though they may now use somewhat different techniques and symbols.

Handbook of Optical Design (MIL HDBK 141). This handbook, which contains much practical design information and very useful discussion of procedures, should soon be published by the Government Printing Office (written 9/10/63). The material was prepared primarily by staff members of the University of Rochester, Institute of Optics. It has also appeared in part as Summer Course Notes. Some copies are available from the Institute.

Herzberger, M., *Modern Geometrical Optics,* Interscience, New York (1958).

Chapter 10

OPTICAL SYSTEMS

Warren J. Smith

Infrared Industries, Inc.

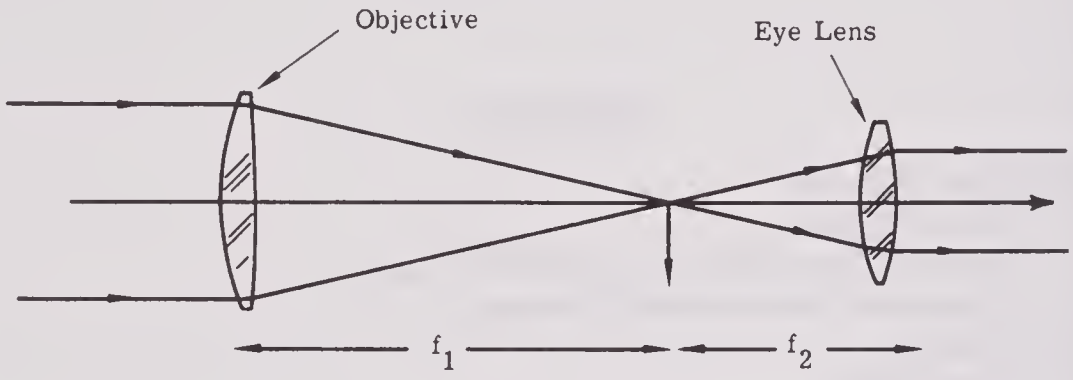
CONTENTS

10.1. Useful Optical Devices 422
10.1.1. Afocal Systems 422
10.1.2. Relay Systems, Erectors, and Periscopes 424
10.1.3. Aplanatic Surfaces and Elements 427
10.2. Detector Optics 427
10.2.1. Application 427
10.2.2. Field Lenses 429
10.2.3. Light Pipes (Cone Channel Condensers) 430
10.2.4. Immersion Lenses 433
10.3. Refracting Objectives 434
10.3.1. The Single-Element Objective 434
10.3.2. Achromats 436
10.3.3. Combinations of Several Thin Single Elements 436
10.4. Reflecting Objectives 437
10.4.1. The Spherical Reflector 437
10.4.2. The Parabolic Reflector 440
10.4.3. Other Conics: The Ellipsoid and Hyperboloid 441
10.4.4. Compound, or Double, Reflecting Systems 441
10.5. Catadioptric Objectives 443
10.5.1. The Schmidt System 443
10.5.2. The Mangin Mirror 444
10.5.3. The Bouwers (Maksutov) Objective 445
10.5.4. Aberrations of a Plane Parallel Plate in a Convergent Beam 447
10.6. Rapid Estimation of Image Blur Size for Several Optical Systems 448
10.6.1. Diffraction-Limited Systems 448
10.6.2. Spherical Reflector 450
10.6.3. Mangin Mirror 452
10.6.4. Parabolic Reflector 452
10.6.5. Schmidt System 453
10.6.6. Catadioptric Systems 453
10.6.7. Refracting Systems 454
10.6.8. Summary of Equations 455

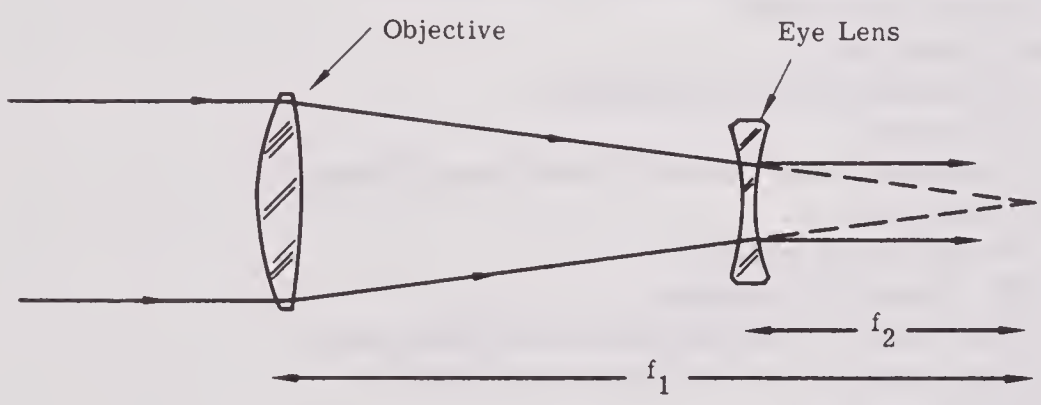
10. Optical Systems

10.1. Useful Optical Devices

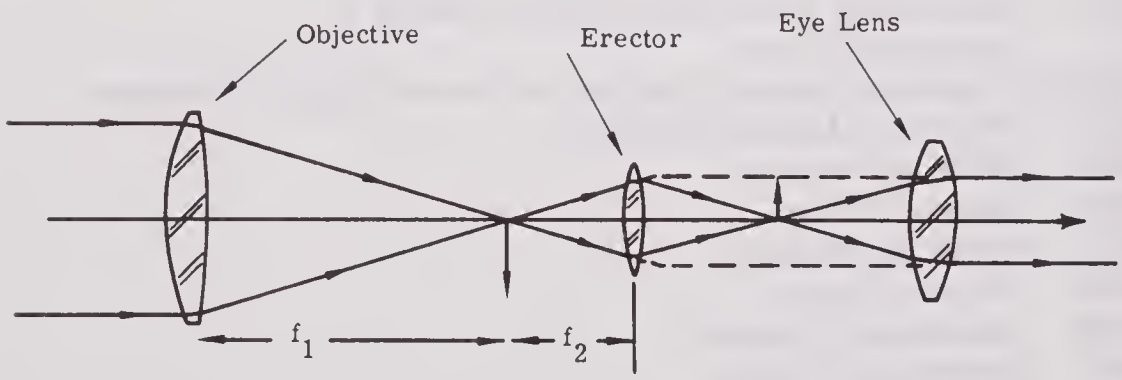
10.1.1. Afocal Systems. An afocal system has its object and image at infinity and thus has no focal length. It is composed of two or more components so arranged that (in a two-component system) the image of the first component, which is the object for the second, lies exactly at the first focal point of the second component and is thus reimaged at infinity. Figures 10-1 and 10-2 show afocal systems.



Astronomical Telescope



Galilean Telescope



Terrestrial Telescope

FIG. 10-1. Astronomical, Galilean, and terrestrial telescopes.

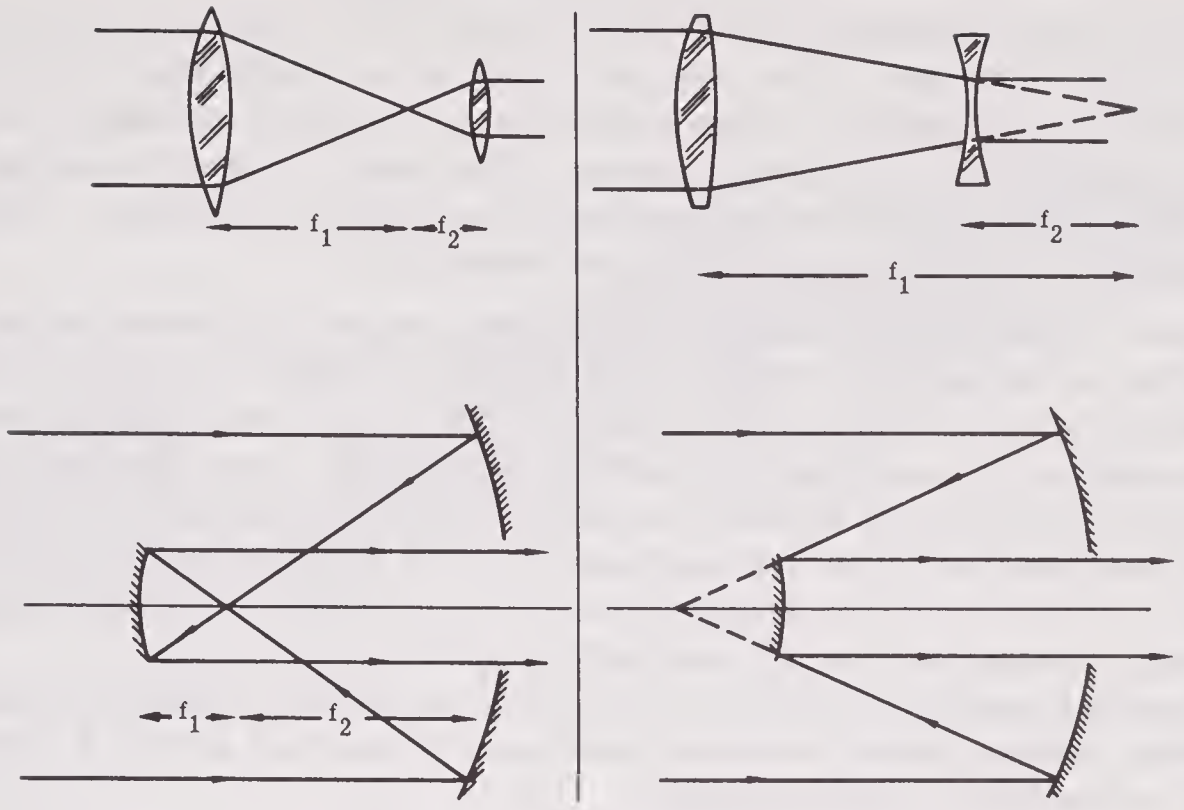


FIG. 10-2. Afocal systems (lower sketches show reflecting equivalents or refracting systems above).

10.1.1.1. *Magnification, or Magnifying Power.* An afocal system usually has a characteristic magnification, in that the object appears larger (or smaller) when viewed through the telescope than with the unaided eye. Figure 10-3 shows a simple telescope with an objective focal length of f_1 and an eye-lens focal length of f_2. The object subtends a half-angle of α_1 at the objective; thus the image formed by the objective has a height equal to $f_1 \tan \alpha_1$. The half-angle subtended by the internal image at the eye lens is thus the image height divided by the focal length of the eye lens, giving

$$\tan \alpha_2 = \frac{f_1}{f_2} \tan \alpha_1 \tag{10-1}$$

Since the apparent angular size of the object when viewed through the telescope is given by $\tan \alpha_2$, the object will appear to have been magnified by f_1/f_2, where the magnification is given by

$$M = f_1/f_2 \tag{10-2}$$

The diameter of the axial bundle of rays emerging from the telescope is also governed by the magnification. From Fig. 10-3,

$$\frac{CA_1}{CA_2} = \frac{f_1}{f_2} = M \tag{10-3}$$

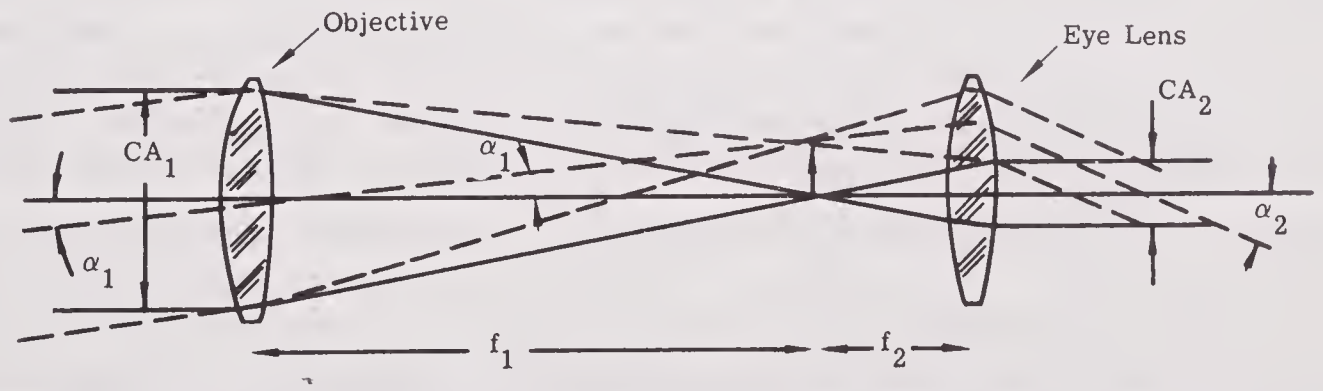


FIG. 10-3. Simple telescope with objective focal length of f_1 and eye lens focal length of f_2.

The space between objective and eye lens is equal to $f_1 + f_2$ for both the astronomical and Galilean telescopes. Also, the eye lenses of the astronomical and terrestrial telescopes form an external image of the objective aperture, through which all the energy emerging from the telescope passes. This image is called the exit pupil and is the customary location of the eye (see Sec. 9.3 on stops and apertures). The Galilean telescope has a virtual exit pupil inside the telescope [1].

10.1.1.2. *Anamorphic Systems.* An anamorphic system has a focal length or magnification in one meridian different from that in the other. In Fig. 10-4 the eye is replaced by a lens of focal length f_3 at the exit pupil. The focal length of the combination is equal to the magnification of the telescope times f_3. If the telescope is composed of lenses with cylindrical surfaces, the overall system will have a focal length of $M \cdot f_3$ in one meridian, but in the meridian parallel to the cylinder axes it will have a focal length of f_3 because in this meridian the cylinders are acting as plane surfaces. A schematic anamorphic system is sketched in Fig. 10-5.

Anamorphic systems can be used in infrared work where it is desired to use a square or circular detector and at the same time obtain a wide field of view in one meridian and a narrow field of view in the other.

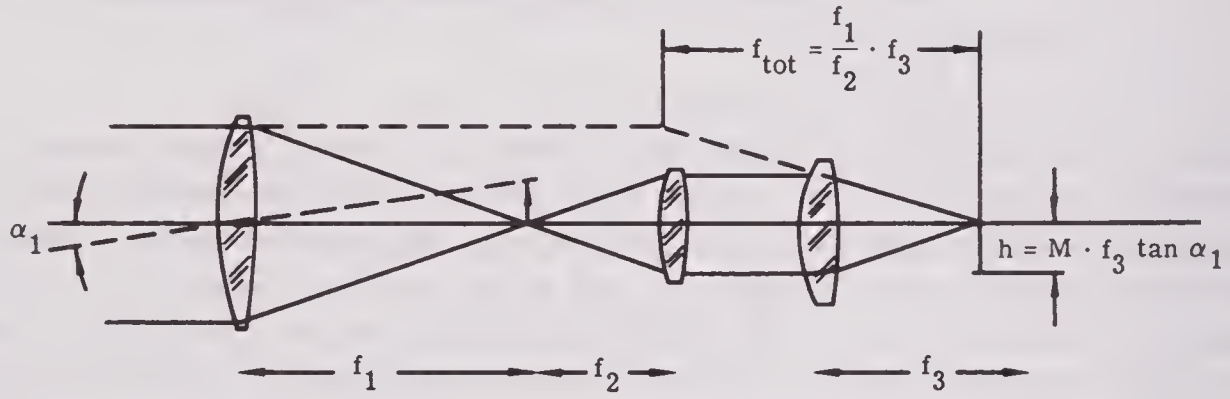


FIG. 10-4. Anamorphic system with eye replaced by lens of focal length f_3 at the exit pupil.

10.1.1.3. *Applications of Afocal Systems.* Figure 10-2 shows the refracting and reflecting equivalents of astronomical and Galilean telescopes. These systems can be used in infrared work to reduce the diameter of the collected bundle of rays from the target. Figure 10-6 shows such an application. Because the field angles are increased by magnification at the same time that the diameter of the bundle of rays is decreased, this device is limited to systems with small fields of view.

10.1.2. Relay Systems, Erectors, and Periscopes. The erector lens of the Terrestrial telescope (Fig. 10-1) indicates the technique of inverting an image with a relay system, which may at the same time magnify or minify the image. To carry an image through a long narrow path, one may use a series of relay systems, with each subsequent system reimaging the original object. Figure 10-7(*a*) illustrates a triple relay system in which each relay lens operates at unit magnification. A field lens can be introduced at (or near) each image plane to redirect the energy from one relay lens into the aperture of the succeeding relay, as shown in Fig. 10-7(*b*). The field lens forms an image of the aperture of one relay system in the aperture of the next.

10.1.2.1. *Projection Systems and Condensers.* A schematic projection system is shown in Fig. 10-8. The reflector is a spherical mirror positioned with its center of curvature at the light source so that it produces an inverted image of the source in the same plane and at the same size as the source. This raises the average effective brightness of the source, either by filling in the gaps or openings in the filament or

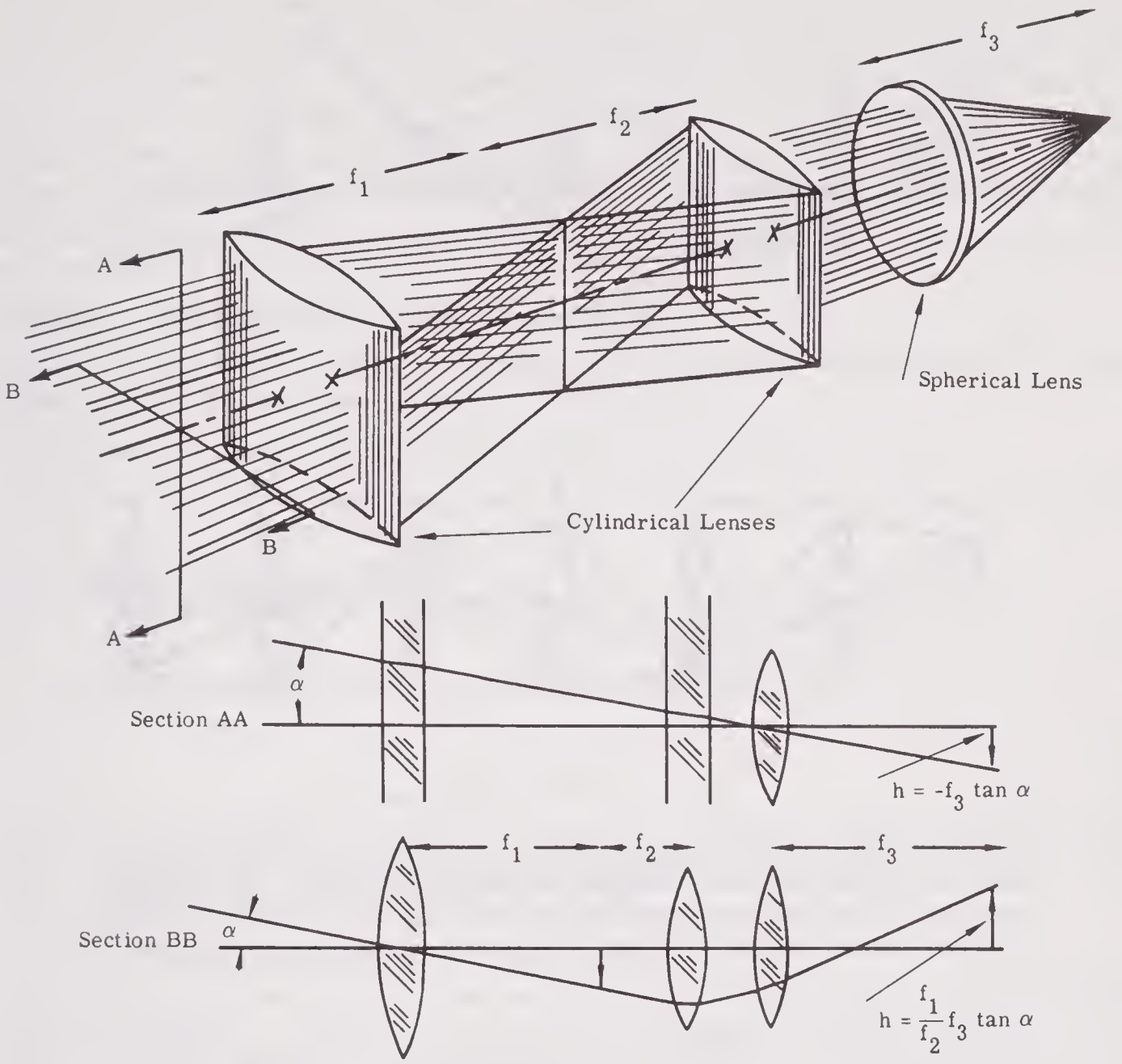


FIG. 10-5. Schematic anamorphic system.

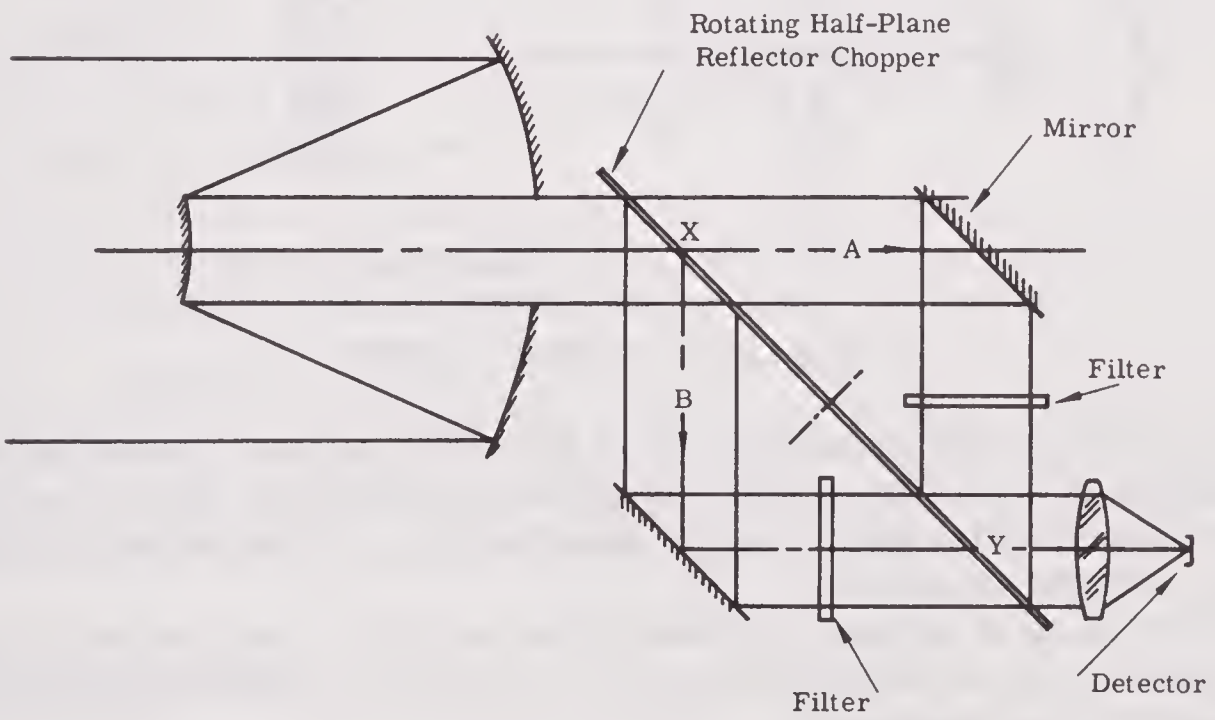


FIG. 10-6. Application of a reflecting afocal system to two-color radiometer. Energy alternately traverses paths *A* and *B*. When the mirrored chopper blade is in the beam at *X*, it is not in the beam at *Y*, and *vice versa*.

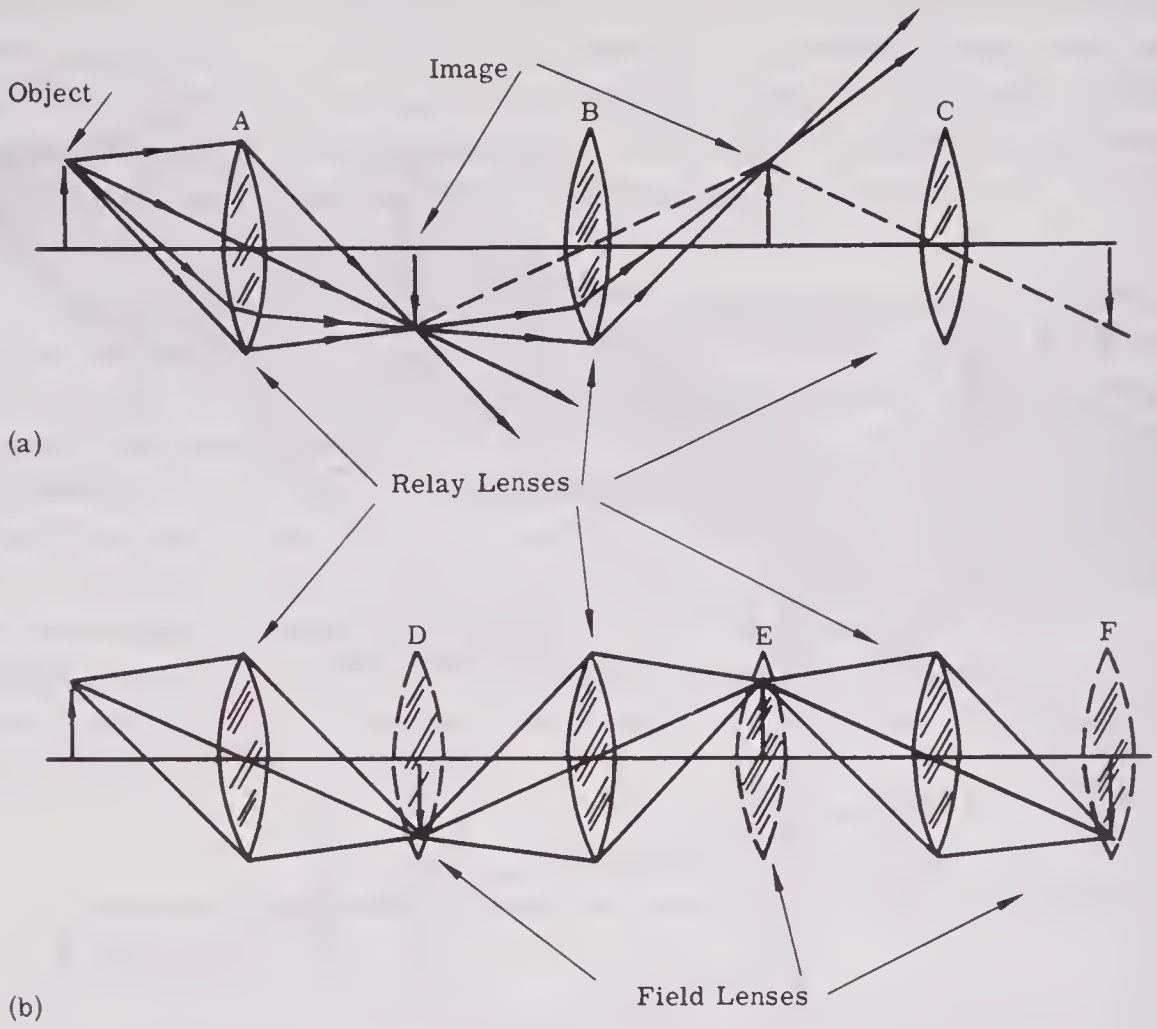


FIG. 10-7. Relay and field lenses: (*a*) A three-element relay system in which most of the light from lens *A* escapes past lens *B* and lens *C*. In (*b*) a field lens *D* is introduced at the image plane to direct light from lens *A* to lens *B*. The field lens is constructed to image the aperture of *A* into the aperture of *B*.

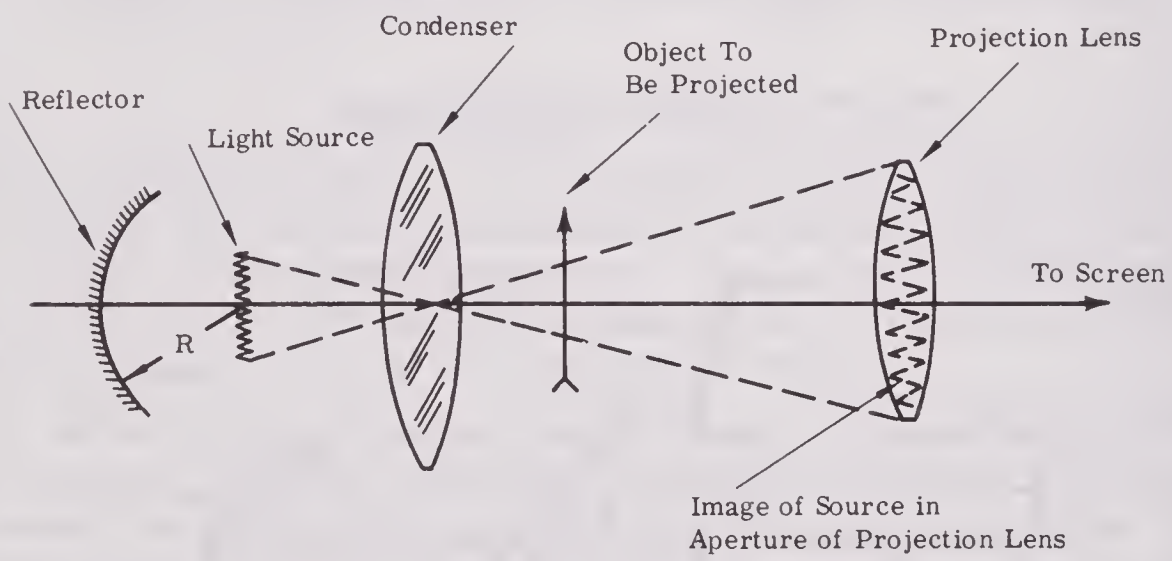


FIG. 10-8. Schematic projection system.

arc, or by raising the temperature slightly, or both. The condenser images the source in the aperture of the projection lens. For optimum efficiency, the image of the source should completely fill the aperture of the projection lens. The projection lens images the object at the desired position.

If the light source is replaced by a detector, the projection system becomes a typical infrared device, with the projection lens, object, and condenser functioning as objective, reticle, and field lens, respectively.

Condensers (and field lenses) are usually made up of one or more single elements. The power and position of the condensing system as a whole are calculated from the

equations given in Sec. 9.2 to produce an image of the desired size and location. An estimate of the number of elements necessary is made, and the total power is divided into that number of parts. Each element is shaped for minimum spherical aberration, either by use of the third-order thin-lens contribution (Eq. 9-215), or by shaping them so that a marginal ray is deviated equally at each surface. The system is then tested, either by trigonometric ray tracing or by constructing a sample. If the residual spherical aberration is too large, the number of elements may be increased, or one or more surfaces may be aspherized.

10.1.3. Aplanatic Surfaces and Elements [2,3]

10.1.3.1. *Aplanatic Surfaces.* The spherical aberration contribution of a surface is zero in the following three cases:

(*a*) When the object and image both lie at the surface ($l = l' = 0$)

(*b*) When the object and image both lie at the center of curvature of the surface ($l = l' = r$)

(*c*) When the object lies at

$$l = r(n' + n)/n \tag{10-4}$$

The following expressions also apply for this case:

$$l' = r(n' + n)/n' \tag{10-5}$$

$$n'/n = \sin u'/\sin u \tag{10-6}$$

Case *c* is the aplanatic condition. A cone of rays may be further converged by the use of an aplanatic surface without introducing any additional aberration. Cases *b* and *c* are both used in detector immersion lenses.

10.1.3.2. *Aplanatic Lenses* [4-7]. A combination of cases *c* and *b* can be used to make a lens which will increase the convergence of a cone of rays without introducing spherical aberration. The first surface is an "aplanatic" surface of the case *c* type which converges the rays by the factor n'/n of Eq. (10-6). The second surface is then made concentric with the image formed by the first surface. The rays are undeviated by the second surface, and the convergence of the entire cone of rays is increased by n'/n, or, if the lens is in air, by its index. Figure 10-9 indicates the layout of such a lens. An aplanatic lens or surface can only increase (or decrease) the convergence (or divergence) of a cone of light. It does not bring parallel rays to a focus nor does it change a diverging cone of rays to a converging one, or *vice versa.*

10.2. Detector Optics

The optics immediately associated with infrared detectors are field lenses, immersion lenses, and light pipes (or cone condensers). All allow use of a smaller size detector. Field lenses and light pipes also redistribute the energy over the surface of the detector.

10.2.1. Application. In Fig. 10-10 the objective is shown as a simple lens; however, the considerations outlined below apply also to a system with a compound objective or a reflecting-type objective. The objective has a focal length f and an effective clear aperture A. The detector size is D and the detector covers a half-angle of view α.

If the objective is well corrected, then the relative aperture is given by

$$f/\text{no} = f/A \tag{10-7}$$

and for moderate angles

$$D = 2\alpha f \tag{10-8}$$

Substituting the value of f given by Eq. (10-8) into Eq. (10-7) gives

$$f/\text{no} = D/2\alpha A \tag{10-9}$$

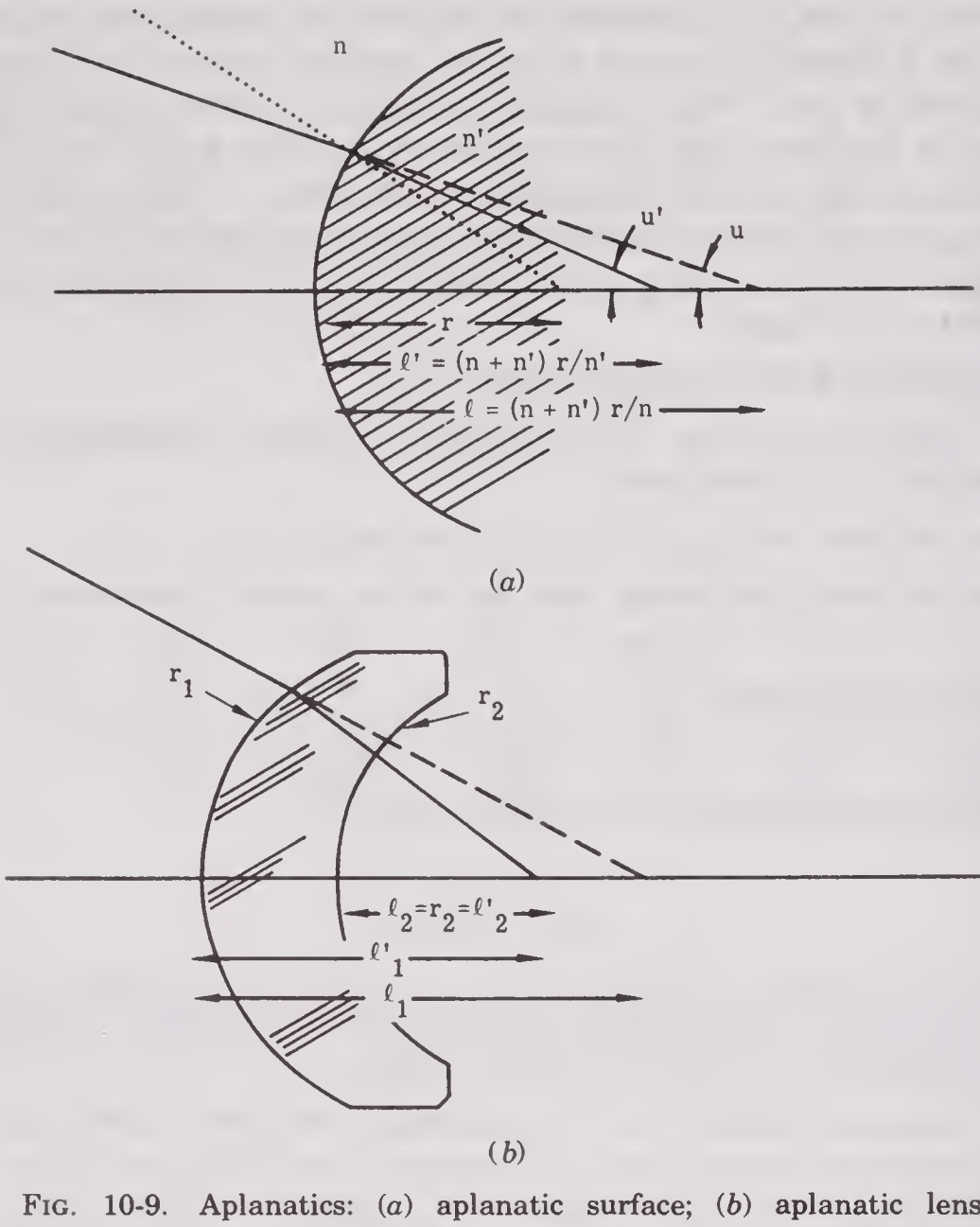


FIG. 10-9. Aplanatics: (*a*) aplanatic surface; (*b*) aplanatic lens where r_1 is an aplanatic surface, r_2 has its center of curvature at l_1'. This lens increases the convergence of the cone of rays by a factor of n without introducing aberration.

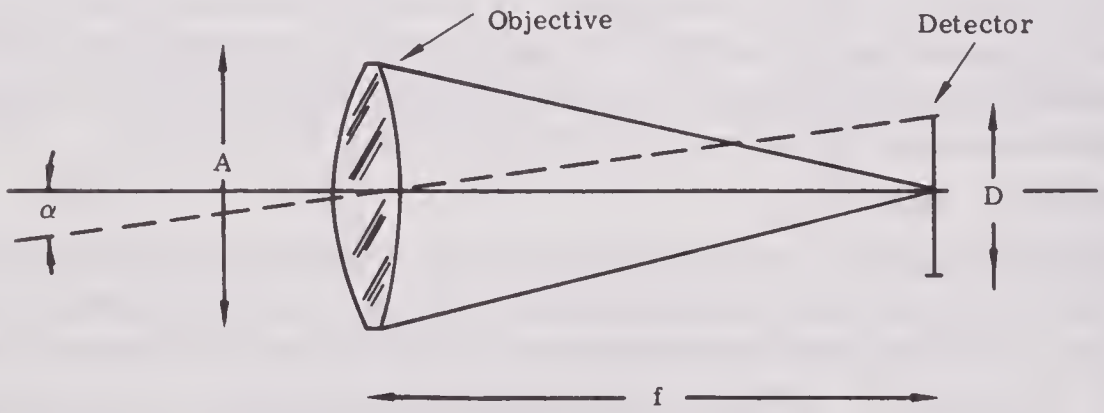


FIG. 10-10. Simple infrared system.

Most infrared systems require a large aperture A to collect a maximum of radiation, a small detector D to optimize the detector signal to noise ratio, and a large field α. These requirements reduce the necessary f/no (that is, increase the "relative aperture"). The limit on the relative aperture of a well-corrected optical system is that it cannot exceed twice the focal length; that is, $f/0.5$ is the smallest f/no attainable (see Fig. 10-11). In a well-corrected system, the Abbé sine condition must hold (p. 373). The sine condition can be expressed as

$$Y = f \sin u' \tag{10-10}$$

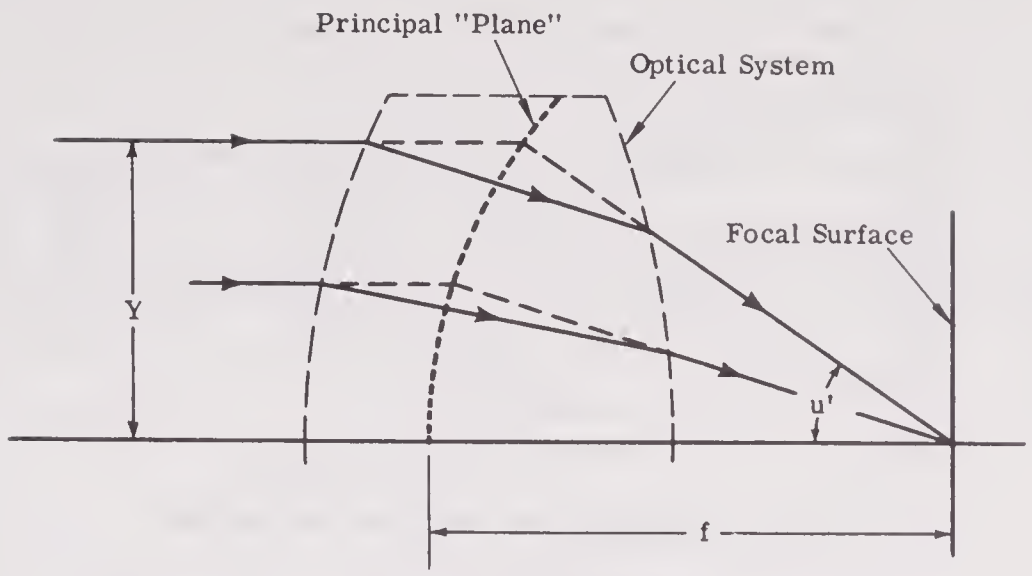


FIG. 10-11. Principal "plane" of well-corrected system is a spherical surface.

The limiting relative aperture is given by

$$f/\text{no} = f/2Y = f/A = 0.5 \tag{10-11}$$

If the f/no from Eq. (10-11) is substituted into Eq. (10-9), the following theoretical limitation on the relationship of A, D, and α are obtained:

$$D_{min} = \alpha A \text{ (theoretical limit)} \tag{10-12}$$

In practice it is often undesirable to exceed a speed of $f/1.0$, in which case the relationship becomes

$$D_{min} = 2\alpha A \text{ ("practical" limit)} \tag{10-13}$$

This limit applies to any optical systems with the detector in air including systems with field lenses and systems with light pipes. Immersion of the detector in a material of index n allows a reduction in D by a factor of $1/n$.

Equations (10-12 and 10-13) are among the first to be considered in designing an infrared system.

10.2.2. Field Lenses. Field lenses are used when a reticle is placed at the focus of the system, illuminating the detector evenly, or when a minimum-size detector is required for use with a primary objective of low relative aperture. Figure 10-12 illustrates the principle of the field lens. The objective forms an image in the focal plane, beyond which the light rays diverge. The field lens causes the rays of light at the edge of the field to be bent toward the axis so that they all fall on the detector. When a field lens is designed, the power and position of the lens are selected so that the objective aperture imaged on the detector is the same size as the detector. Equations (9-2) and (9-4) are used to determine the relationships between the quantities of Fig. 10-12:

$$A/D = S/S' \tag{10-14}$$

$$\frac{1}{f} = \frac{1}{(-S)} + \frac{1}{S'} \tag{10-15}$$

Combining and solving for the focal length of the field lens gives

$$f = \frac{SD}{(D + A)} \tag{10-16}$$

The necessary diameter of the field lens can be deduced from the path of the extreme off-axis rays (shown dashed in Fig. 10-12). Vignetting these rays at the edge of the

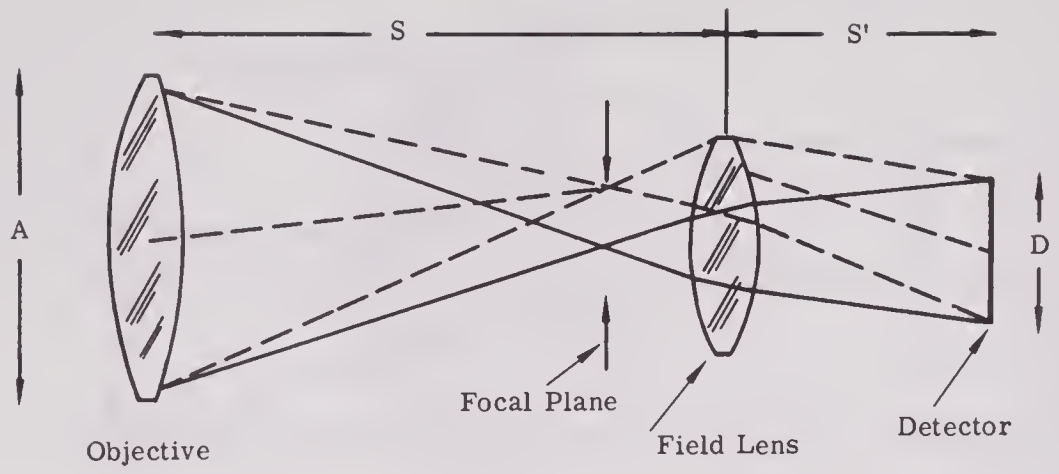


FIG. 10-12. Function of field lens in infrared system.

field reduces the diameter of the field lens. The optimum location for the field lens is at the focal plane, where its diameter is minimized; practical considerations of interference with the reticle, or of the effects of small imperfections in the field lens upon system performance, usually enforce a compromise location. The illumination at the detector under these conditions is now identical to that at the aperture of the objective. If there is no vignetting, a point image may be moved all over the field of view without affecting the intensity or distribution of the illumination at the detector.

Usually the objective aperture is circular and the detector is square. If the detector's signal-to-noise ratio varies inversely with the square root of its area, or directly as $(1/D_1)$, and if the optical limit of Eq. (10-11 and 10-12) has been attained with the image of the objective aperture forming a circle of illumination inscribed within the square detector, the detector can be reduced in size. A loss of only 10% in S/N occurs if the detector is reduced in size until its square is inscribed in the illumination circle.

The field lens is designed to minimize aberrations and place the maximum amount of energy on the detector. The technique is to consider the field lens as an image-forming system in itself, and to evaluate (by ray tracing or third-order contributions) the quality of the image of the objective aperture which it forms. When high relative apertures are required of the field lens, the requirements for correction of coma, and to a lesser extent spherical aberration, usually lead to a two- (or more) element system, as shown in Fig. 10-13. The first element is often planoconvex and the second a meniscus in a shape to minimize aberration and optimize the energy delivered to the detector.

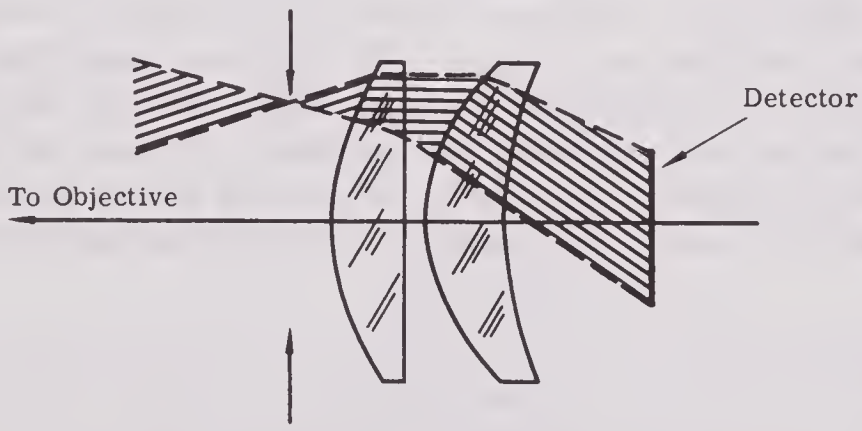


FIG. 10-13. Section through typical field lens used with infrared detectors.

10.2.3. Light Pipes (Cone Channel Condensers) [8,9]. A light pipe is a hollow (or solid) truncated cone (or pyramid), with highly reflecting walls, which collects light at its receiving end and channels it by successive reflections to the other end,

at which a detector is usually located. It serves the same function as a field lens and is subject to the same optical limitations on the relationship among system aperture, detector size, and system field of view (Eq. 10-7 through 10-12). A completely reflecting (hollow) light pipe can be constructed for situations where there is no suitable refracting material. Such a light pipe "scrambles" the imagery between detector and objective aperture.

Figure 10-14 shows a section of a typical light pipe and the path of a ray through the pipe. Ordinarily the detector is located at the smaller end of the pipe and the image plane of the system is located near the large end.

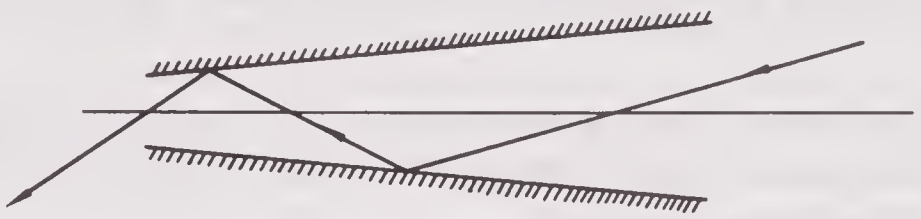


FIG. 10-14. Section through light pipe, with path of ray showing pipe function.

Figure 10-15 shows the construction used for tracing the path of a ray through a light pipe. S_1 and S_2 are the upper and lower sides of the pipe, respectively. Since the image from a plane reflecting surface is located an equal distance behind the surface, the image of S_2 in S_1 is a line through the intersection of S_1 and S_2, making the same angle (α) to S_1 that S_2 does. Subsequent reflected images of the sides are similarly constructed. Then a meridional ray through the system can be represented by a straight line, and the intersections of the line with the images of the surfaces are the locations of the reflections of the ray. In Fig. 10-15 the ray AB, which is shown by the straight line AB', emerges from the pipe after three reflections. The ray represented by AC, however, never reaches the small end of the pipe. It is turned around and eventually emerges from the large end of the pipe. The limiting case is a ray tangent to circle FG, which is the locus of the reflected images of the exit window of the pipe.

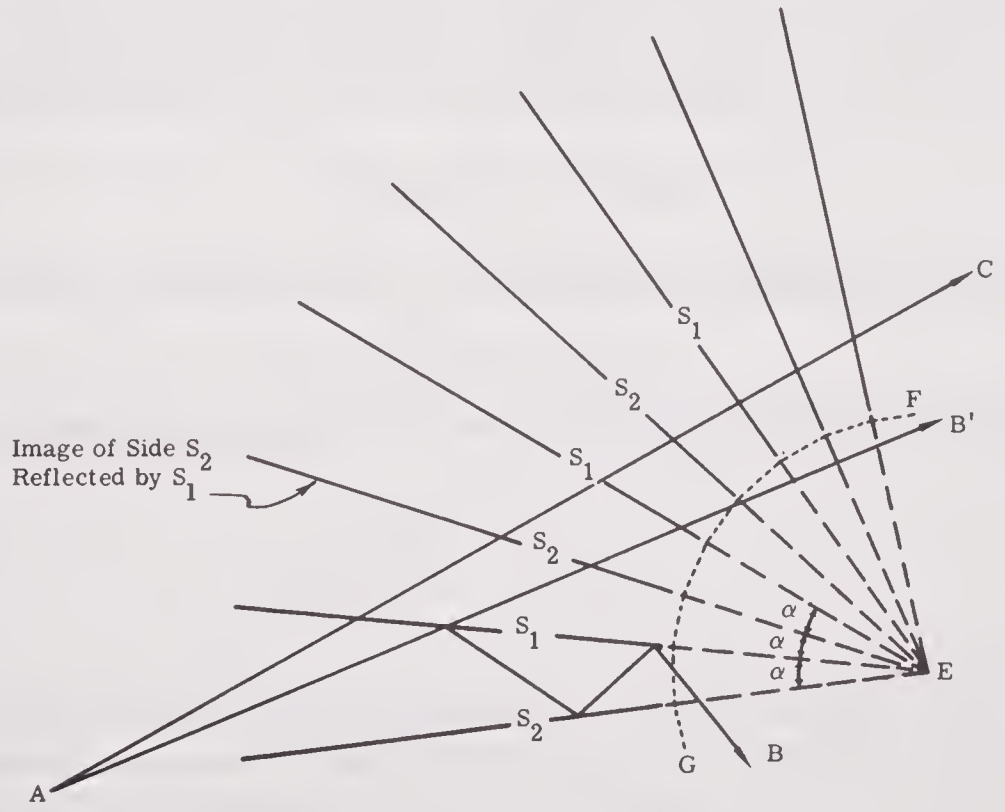


FIG. 10-15. Graphic ray tracing through light pipe.

If the large end of the pipe represents the field stop of an objective system and the small end the detector size, the minimum detector size is given by

$$D_{min} = \alpha A \tag{10-17}$$

If x is the length of the light pipe, c the radius of the detector, and s the radius of the large end of the pipe,

$$x = 1 - \frac{c}{s}\frac{s\cos\alpha}{(c/s) - \sin\alpha}$$

There are a number of limitations on solid light pipes. If the sides of the pipe are made reflecting and the detector is in optical contact with the pipe (and of the same or higher index), the situation is much the same as that described above except that the rays are refracted toward the pipe axis upon passing through the entrance face. This increases the acceptance angle of the pipe by a factor equal to the index n, and the limiting size of the detector becomes

$$D_{min} = \alpha A/n \tag{10-18}$$

Some solid pipes are made without any reflecting coating, and operate on the basis of total internal reflection. Total internal reflection (T.I.R.) occurs when light passes from a medium of higher index to a medium of lower index at an angle of incidence such that the angle of refraction is 90°.

Figure 10-15 shows that the angle of incidence decreases as the ray passes down the pipe, thus establishing a limiting length on the pipe under these conditions. A construction similar to that of Fig. 10-15 can be made, except that at each surface the critical angle must not be exceeded. This limits the maximum acceptance angle to

$$\arcsin\left[(n^{1/n})^2 - 1\right]^{1/2}$$

Another limit for solid pipes occurs if the detector is not in contact with the end of the pipe, or if the detector or its coverplate have a lower index than the pipe. Under these circumstances T.I.R. can occur at the end face of the pipe and prevent the light from leaving the pipe.

The transmission efficiency of a light pipe is a function of the reflectivity of the surfaces, and if the pipe is solid, of the transmission of the material used. For meridional rays (as in Fig. 10-15) it is simple to determine the path length through the medium and the number of reflections for any given ray. The path length d is given by

$$d = L\left(1 - \frac{n_1}{n}\sin^2\theta\right)^{-1/2} \tag{10-19}$$

where θ = angle of incidence on front surface. The transmission is given by

$$\tau = \rho^{m/2}(1 - \rho^2)\, e^{-\alpha d} \tag{10-20}$$

where ρ = measured reflectance of the sides of the pipe; m, the number of reflections, is given by

$$m = \frac{nL\sin\theta}{2r(n'^2 - n^2\sin^2\theta)^{1/2}}$$

and r = radius of pipe.

Light pipes are made in the form of truncated cones, cylinders, and pyramids (commonly four sided). Figure 10-15 shows that from point A one sees a mosaic of reflected images of the detector laid on a spherical surface about point E. The light is "scrambled" or "folded" over the detector surface and, if there is a reflection loss at the surfaces,

provides a ready method of visualizing the decreased effective sensitivity at the edges of the detector mosaic. When the light pipe is conical, the analysis is more complex, because the first-order reflected images of the detector will blend into a single annular image, and each of the higher-order images forms its own annulus.

When a light pipe is designed, convenient length and a detector size within the limits of Eq. (10-12) and (10-18) are chosen. A drawing of the system, in the manner of Fig. 10-15, is made, and a set of "typical" rays (usually the full aperture rays to an on-axis point and to a point at the extreme edge of the field) are drawn in. These rays are examined and adjusted for the limitations described above. The longer the light pipe, the greater the number of reflections and the greater the transmission path if the pipe is solid.

10.2.4. Immersion Lenses. The immersion lens is used to decrease the detector size to obtain increased detector signal-to-noise ratios. The aplanatic surface described in Sec. 10.1.3.1. is used for this purpose.

10.2.4.1. *The Hemispherical Immersion Lens.* The immersion lens with greatest applicability is the hemispherical element with the detector located at the center of curvature (see Fig. 10-16). In an optical system with the detector located either at the focus of the primary system (objective) or at the image of the objective aperture formed by a field lens, the energy focused on the detector converges with a half-angle u and the detector size corresponds to h, the maximum height of the image. Since this image is formed in air of index 1.0, the Lagrangian invariant is hu (Eq. 9-6). If a hemispherical lens is inserted in the system with its center of curvature at the focus of the rays, the convergence angle is not changed (since the rays of the axial cone strike the surface normally), and the Lagrangian invariant is

$$h'\ n'\ u', \text{ which then is the same as } h'n'u \qquad (10\text{-}21)$$

Since the invariant for the image in air must equal that for the image in the immersion lens, the two may be equated to yield

$$h' = h/n' \qquad (10\text{-}22)$$

This usage of the hemispherical surface corresponds to case *b* of Sec. 10.1.3.1.

Design. To design an immersion lens, a diameter for the lens is selected larger than the unimmersed detector size and a few rays are traced from the primary optical system directed toward the edge of the field. The focus of these rays is examined to locate the optimum position for the detector. If the angles at which the rays strike the surface of the immersion element are large, the loss due to reflection may be too great. These angles may be reduced by increasing the radius of the immersion element. The radius used for an immersion element is usually a compromise among the desire for reasonable incidence angles, the space available for the element, and the relatively high costs of the material for immersion lenses.

Immersion. To obtain the fullest gain from an immersion lens, the detector must be in optical contact with the plane surface. Some detectors can be deposited directly on this surface if the lens is made of a material compatible with the electrical and mechanical characteristics of the detector. If the detector index is higher than that of the immersion lens, the situation is ideal. If the detector index happens to be lower than that of the lens, or if there is a layer of air, or cement, or a low-index cover plate between the lens and the detector, then total internal reflection can occur at the interface where the light travels from high index to low. Since high-efficiency systems must operate with high convergence angles at the detector, the angles of incidence at these interfaces are often very large and T.I.R. is a problem unless the detector is truly immersed.

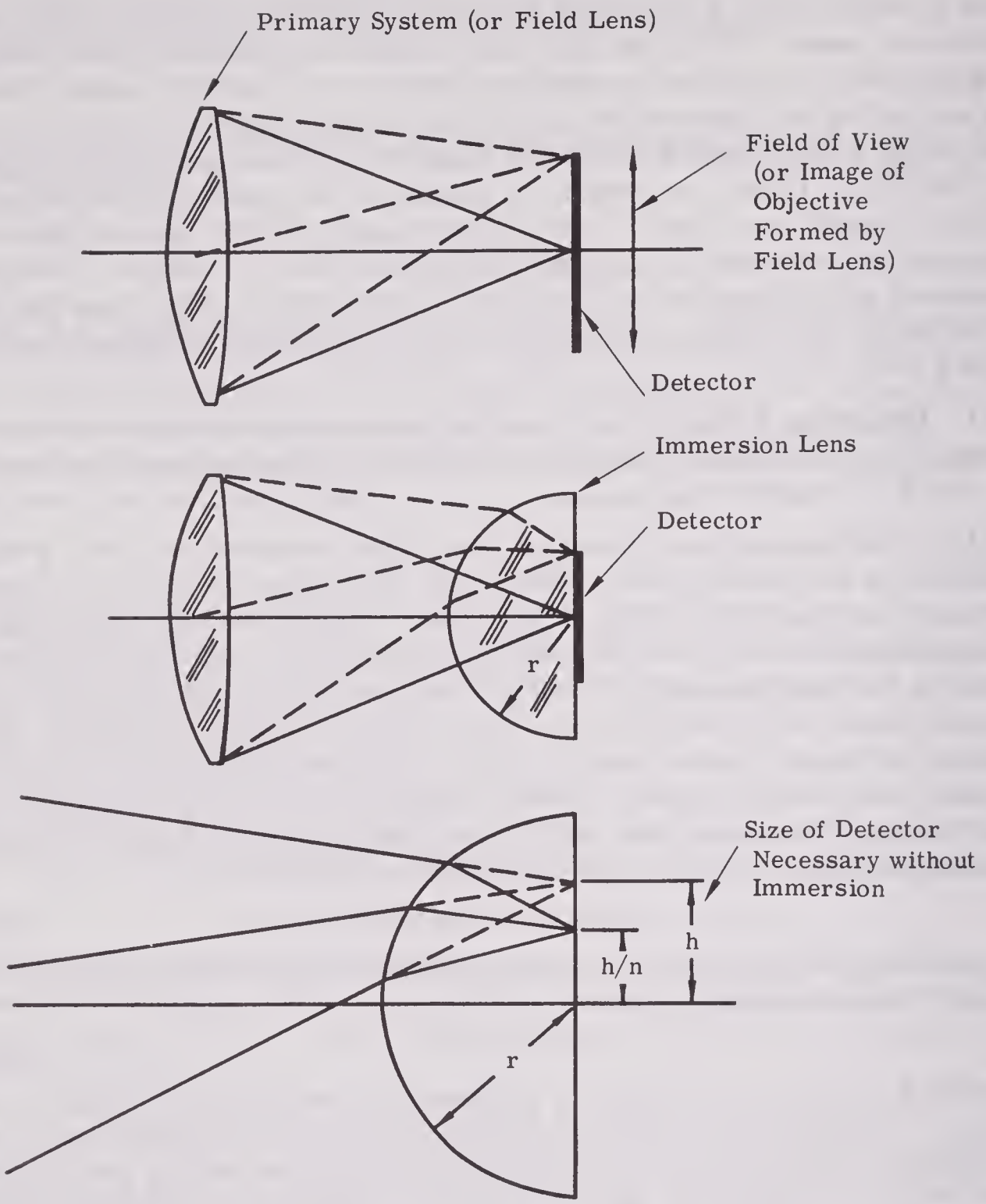


FIG. 10-16. Function of hemispherical immersion lens in reducing the size of the detector necessary to cover a given aperture or field of view.

10.2.4.2. *The "Aplanatic" Immersion Lens.* Case *c* of paragraph 10 1.3.1. provides another possible surface for use in an immersion lens, since the "aplanatic" surface may be inserted into a convergent beam without introducing aberration, just like the concentric surface of the hemispherical lens. As shown in Fig. 10-17, the use of this type of lens is limited to low-aperture systems.

10.3. Refracting Objectives

10.3.1. The Single-Element Objective. A single refracting element usually forms a very imperfect image. Chromatic aberration spreads out the image, and unless expensive aspheric surfaces are used, there is spherical aberration. Because the image quality required of many infrared systems is relatively poor, the simplicity and economy of a single spherical lens make it useful.

The image quality of a single element can be estimated from the thin-lens contribution equations of Sec. 9.6.3. (Figures 10-35 through 10-41, except Fig. 10-36, allow

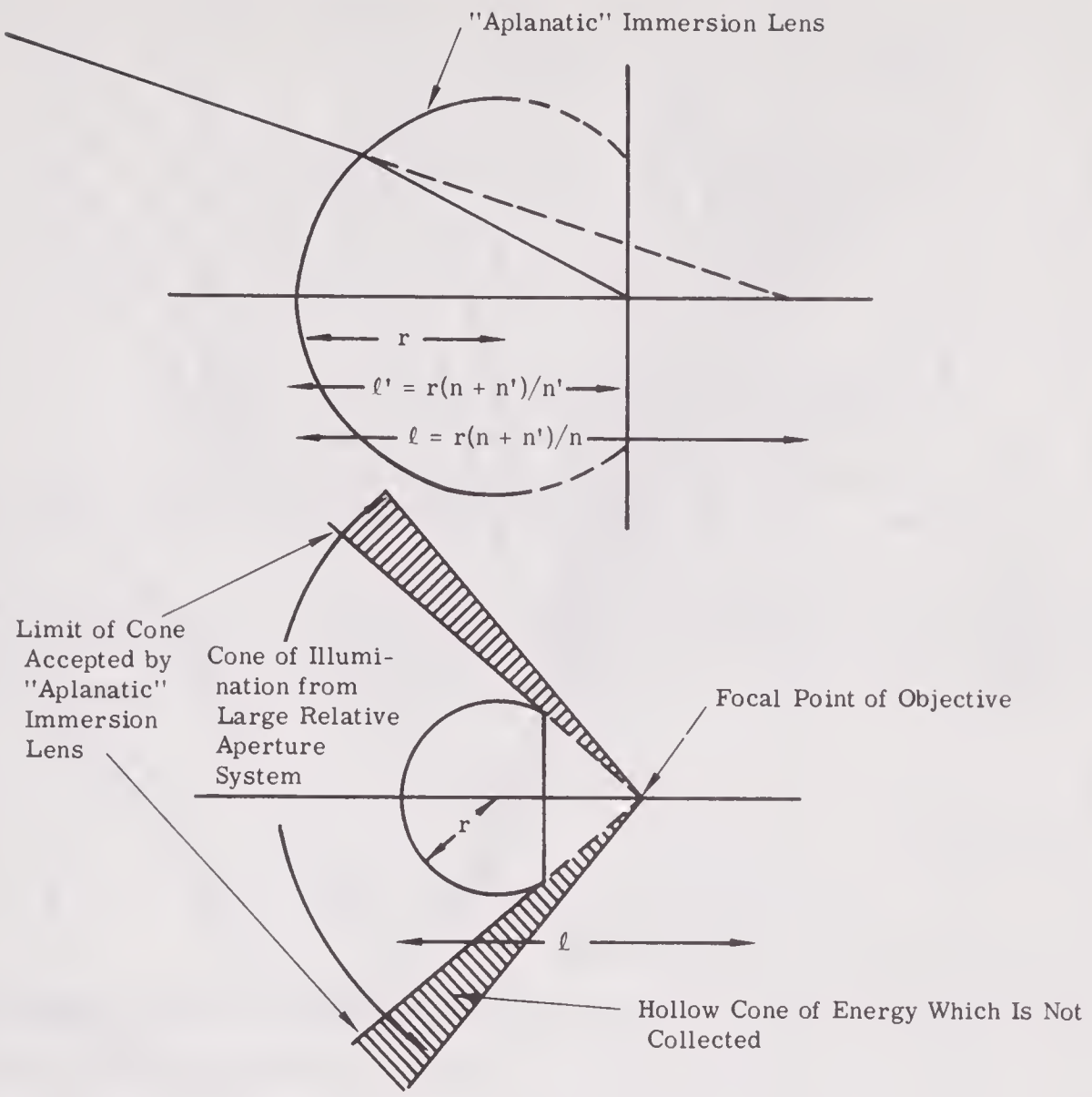


FIG. 10-17. Construction of "aplanatic" immersion lens showing how the acceptance angle is severely limited even for a zero angle field of view system, shown for $n = 3.5$ and an $f/0.8$ beam.

estimation of the angular blur produced by a single element.) It is possible to determine the size of the blur spot due to spherical aberration from Fig. 10-18. Each line on the plot represents an index of refraction. The x axis is the shape of the element, and the y axis is the size of the angular blur spot. To determine the spherical blur, one locates the proper refractive index line and finds the point on the line corresponding to the shape factor (K) of the lens. The shape factor is given by

$$K = c_1/(c_1 - c_2) \tag{10-23}$$

where c_1 and c_2 = the curvatures (or reciprocal radii) of the surfaces of the lenses. The blur factor is then read from the y axis. To convert this to angular blur in radians, multiply by $y^3\phi^3$, where y is the semi-aperture of the lens and ϕ is the power or reciprocal focal length. The chart is derived from the equation for the third-order spherical aberration of a thin lens presented in Sec. 9.6.2, and applies to lenses with the object at infinity. The chart can be used for finite conjugates by splitting the lens into two portions (each of which is working at infinity) and adding the angular blur of both parts to get the final blur.

To estimate chromatic blur, the index of refraction of the lens material is determined for the extreme wavelengths of the spectral range over which the system is to be used. Divide their difference ($n_v - n_r$) into the median index minus one to get the V value:

$$V = (n - 1)/(n_v - n_r) \tag{10-24}$$

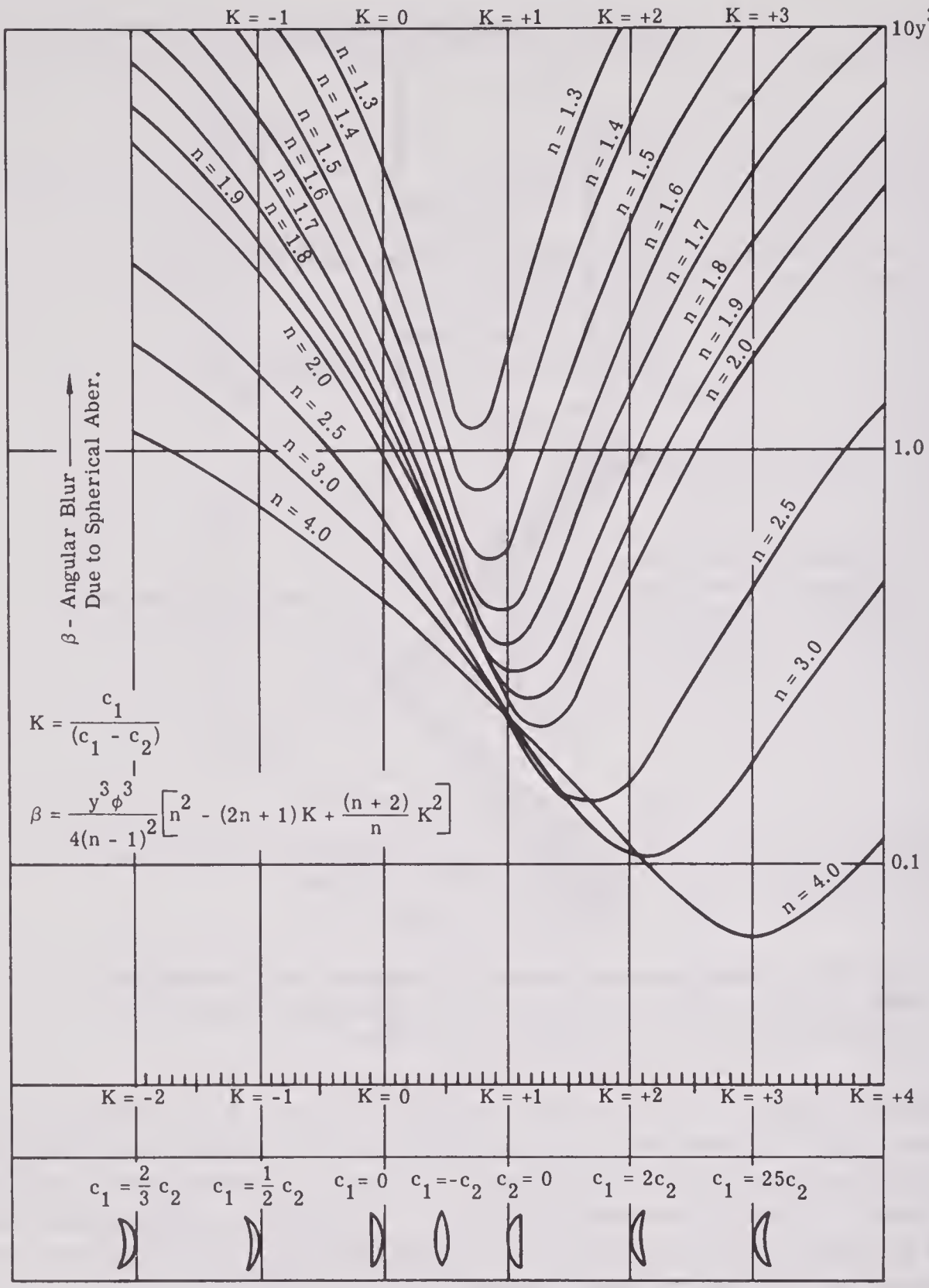


FIG. 10-18. Angular blur of simple lens as function of shape for various indices of refraction.

Then for a lens at infinity the longitudinal chromatic aberration is

$$LchC = f/v \tag{10-25}$$

and the angular chromatic blur is

$$\beta ch = \frac{LchC}{2(f/no)f} = \frac{1}{2V(f/no)} \tag{10-26}$$

10.3.2. Achromats. Refer to Sec. 9.8.

10.3.3. Combinations of Several Thin Single Elements. The use of a series of thin elements as shown in Fig. 10-19 makes it possible to reduce the spherical aberration of a refracting lens, although the chromatic aberration remains the same as for a single element.

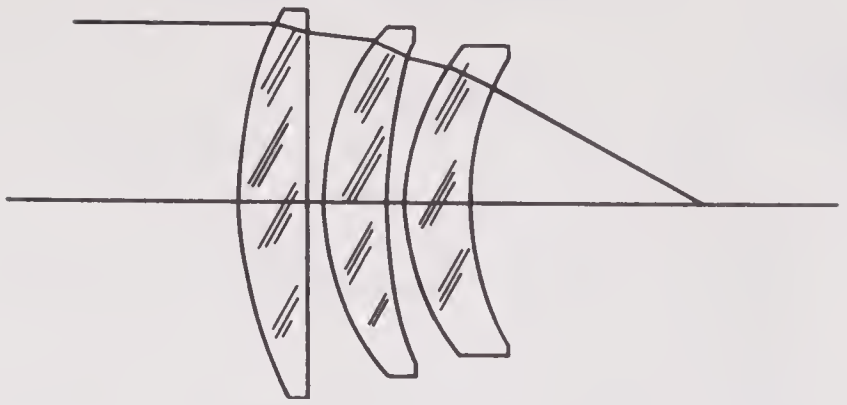


FIG. 10-19. Series of thin spherical lenses, each shaped to minimize spherical aberration. At an index of 1.75 this three-element system can be designed to be completely free of spherical aberration.

If a set of elements with object at infinity is to have a total power ϕ, and each element of the set is bent to the shape for minimum spherical aberration, then the longitudinal spherical aberration of the set is given by

$$SC = \sum_{j=1}^{j=i} SC_j \tag{10-27}$$

$$SC_j = \frac{-Y^2\phi_i n\,[4n-1-4j(j-1)(n-1)^2]}{8i^3(n-1)^2(n+2)} \tag{10-28}$$

where i = the number of elements in the set

Y = the ray height at all elements

ϕ_i = the power of each element

n = the index of the elements

j = the element number

Equation (10-28) has been worked out for $i = 1, 2, 3$, and 4 and is presented in Fig. 10-20. To use the graph, the number of elements i and the index n are determined. The intersection of the i line and the n coordinate extended to the y axis indicates either the longitudinal spherical or the angular diameter of the blur spot.

To fabricate such a system the radii must then be selected so that each element is bent to minimize its spherical aberration contribution. This can be done with Eq. (9-215) and also Eq. (9-203), (9-204), (9-205), (9-206), and (9-226).

It is practical to use these systems at finite conjugates by arranging two such systems (of appropriate focal lengths) facing each other with parallel light between them, as shown in Fig. 10-21.

10.4. Reflecting Objectives

The pure reflecting objective is completely free of chromatic aberration and has no transmission losses. A reflector surface is usually coated with a thin evaporated aluminum film, which has an excellent infrared reflectance and which in itself is durable. Other materials are occasionally used for reflecting films.

10.4.1. The Spherical Reflector. The simple spherical reflector is a useful infrared device. It is simple, inexpensive, and easy to align and mount. When used with an aperture stop at its center of curvature, it has only spherical aberration. Figure 10-22 illustrates this usage of a spherical mirror. When the stop is at the center of

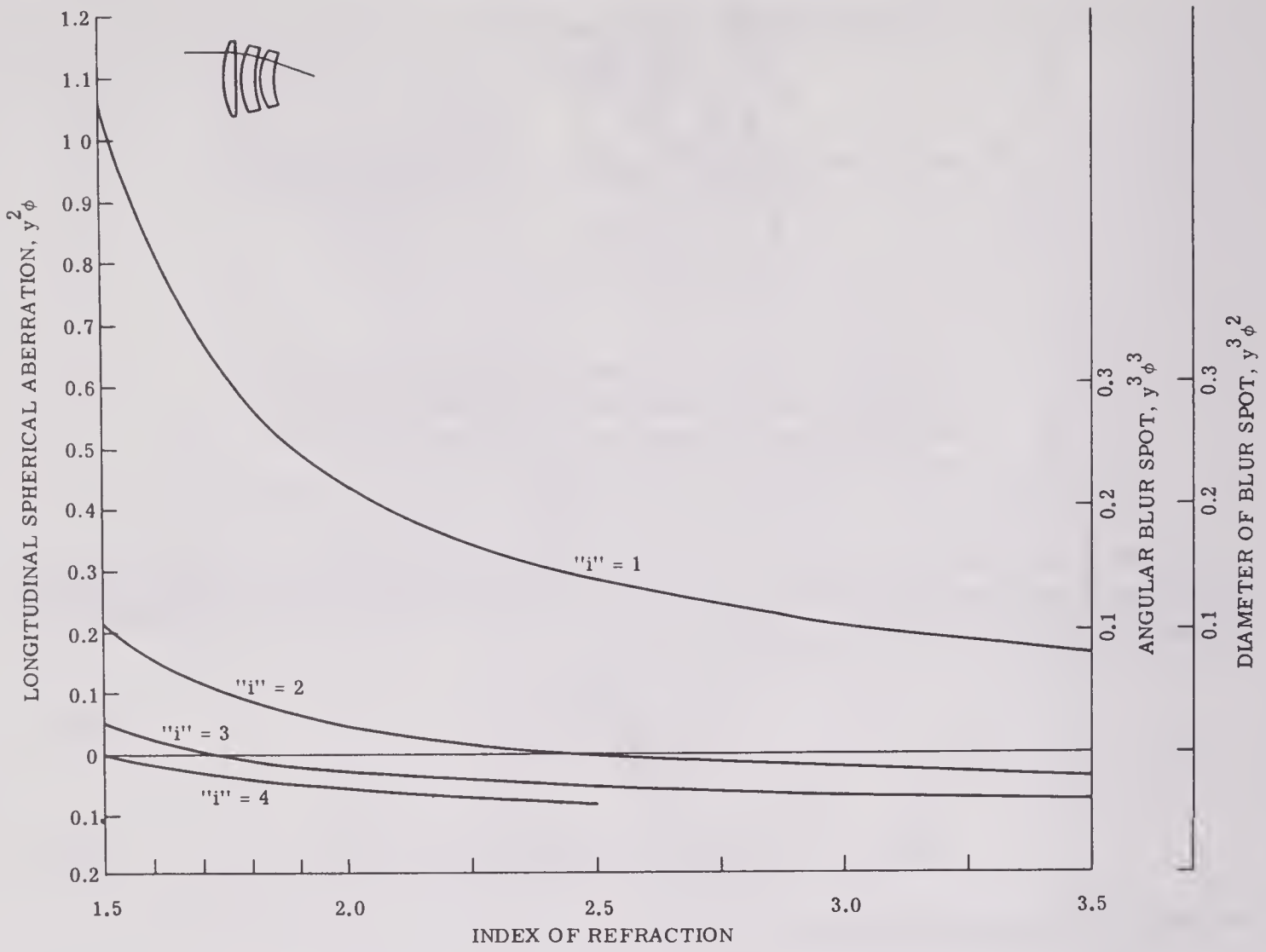


FIG. 10-20. Thin lens aberration of i thin elements of index n, each element bent for minimum spherical aberration.

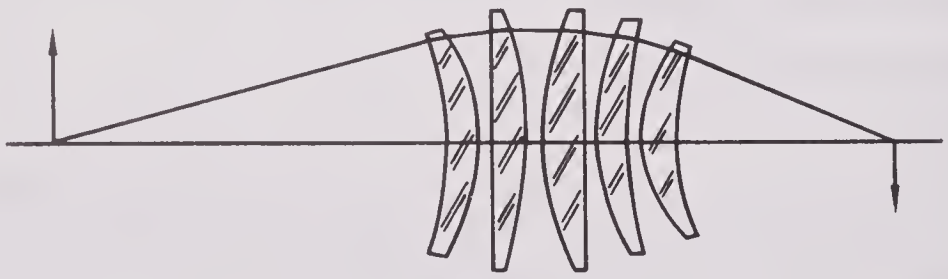


FIG. 10-21. Minimum spherical elements as relay system. For smaller fields, the image quality of this type of system is limited only by chromatic aberration, since spherical aberration is eliminated and coma is very small.

curvature, any principal ray (through the center of the stop) may be considered an axis of the system. Thus the image quality for any field of view is almost the same as the on-axis image quality. The image surface is a sphere with a radius approximately one-half of the mirror radius and concentric with the mirror.

The aberrations of a spherical mirror can be estimated by the third-order surface-contribution equations of Sec. 9.6.2. By setting $n = 1.0$ and $n' = -1.0$, the following expressions can be obtained for the third-order aberrations of a sphere with object at infinity:

Spherical:

$$SC = y^2/4r \tag{10-29}$$

Sagittal coma:

$$CC^* = y^2(l_p - r)u_p^2/2r^2 \tag{10-30}$$

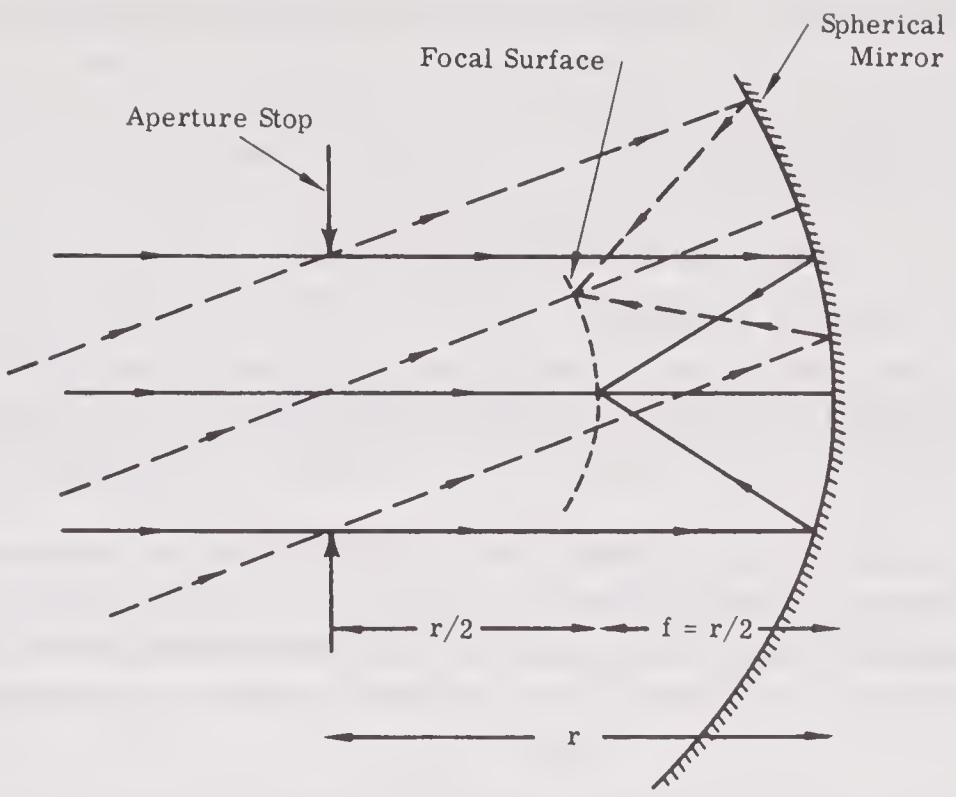


FIG. 10-22. Spherical reflector with stop at center of curvature. In this arrangement the mirror has no coma or astigmatism and the focal surface is a sphere of radius $r/2$ about the center of curvature. Only spherical aberration limits the image quality.

Astigmatism:

$$AC^* = (l_p - r)^2 u_p{}^2/4r \tag{10-31}$$

Petzval curvature:

$$PC = u_p{}^2 r/4 \tag{10-32}$$

where y = the height of the axial ray from the axis

r = the radius of the sphere

l_p = the distance from sphere to the stop

u_p = the slope of the principal ray

$*$ = new third-order coefficient

Note that, if the stop is placed at the center of curvature, $r=l_p$, $AC^*=0$. and $CC^* = 0$. If the stop is placed at the mirror, however, $l_p = 0$ and Eq. (10-30) and (10-31) reduce to

Sagittal coma:

$$CC = -y^2 u_p/2r \tag{10-33}$$

Astigmatism:

$$AC = -r u_p{}^2/4 \tag{10-34}$$

The sagittal and tangential field curvature can be obtained from

$$X_s = PC + AC = \frac{u_p{}^2 r}{2} + \frac{u_p{}^2 l_p}{4}\left(\frac{l_p}{r} - 2\right) \tag{10-35}$$

$$= \frac{u_p{}^2 r}{2} \quad \text{if} \quad l_p = 0$$

$$X_t = PC + 3AC = u_p{}^2 r + \frac{3}{4} u_p{}^2 l_p \left(\frac{l_p}{r} - 2\right) \tag{10-36}$$

$$= u_p{}^2 r \quad \text{if} \quad l_p = 0$$

Where the sphere is used at finite conjugates the following expression is useful:

$$SC = (m-1)^2 y^2/4r \tag{10-37}$$

where m is the magnification,

$$m = \frac{h'}{h} = \frac{l'}{l} = \frac{u}{u'}$$

10.4.2. The Parabolic Reflector. For distant objects located on the optical axis, a paraboloid is completely free of aberrations and its image quality is diffraction limited. The axial image of a perfectly made parabola is an airy disc as described in Sec. 9.9. The paraboloid is used in many infrared systems requiring good image quality over a very small field.

Off axis the paraboloid has the same amount of coma and astigmatism as a spherical mirror when the stop is at the mirror (in both cases). The stop shift theory of Sec. 9.6.3 indicates that, because there is no spherical aberration, coma will not change as the stop is moved. Thus for all stop positions the coma of a paraboloid is given by

$$CC^* = y^2 u_p/4f \tag{10-38}$$

Since a paraboloid has coma, the astigmatism will vary with stop position. The astigmatism with the stop at the mirror is given by

$$AC = -fu_p^2/2 \tag{10-39}$$

and drops to zero when the stop is at the focal plane position ($l_p = f$). The Petzval surface is the same as that of a sphere. Thus with the aperture at the focal plane, both s and t foci lie on a spherical surface of radius f.

A high-quality paraboloid may be an order of magnitude more costly than that of an equivalent spherical surface, which is often of much better optical accuracy. Deep paraboloids are especially difficult to fabricate precisely. Before specifying a paraboloid for a system, one must investigate whether a spherical surface will meet the requirements of the system.

The optical collimator produces a source of radiation which appears to be at infinity, so that rays from any given point of the source are parallel to each other. The prime purpose for most collimators is the testing of other optical systems, and thus their optical quality must be superior to the device being tested. Since a paraboloid has no aberrations (when used on axis), it is an ideal objective for use in an infrared collimator. To avoid obscuration of the central portion of the collimated beam by the light source, the aperture of the collimator is often decentered from the paraboloid axis, resulting in the so-called off-axis paraboloid as illustrated in Fig. 10-23.

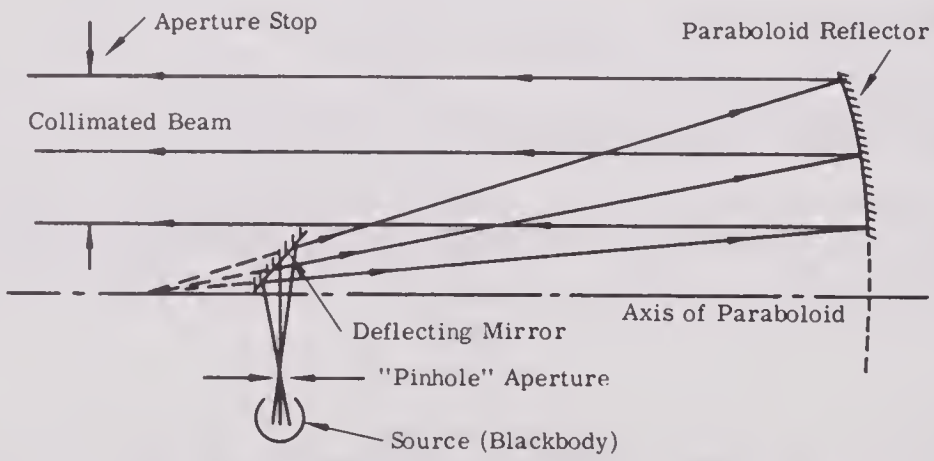


FIG. 10-23. Typical infrared collimator with paraboloid reflector objective, showing use of off-axis aperture.

10.4.3. Other Conics: The Ellipsoid and Hyperboloid. In the ellipsoid and the hyperboloid, rays directed toward, or emerging from, one focus, are reflected toward the other without spherical aberration. Thus the sphere has perfect imagery of points at its center, which is the locus of both foci. At the other extreme, the paraboloid has perfect imagery between a focus at infinity and the point usually thought of as its focus. The ellipsoid and the hyperboloid image from one focus to the other. Both, however, have large amounts of coma so that they cannot be used to form good images of extended objects.

The fabrication costs discussed in Sec. 10.4.2. are even more applicable to the ellipsoid and hyperboloid, as is the advisability of considering the substitution of a spherical surface whenever possible.

10.4.4. Compound, or Double, Reflecting Systems. In many applications the location of the image plane in the path of the incoming rays is inconvenient. Thus a great number of objective systems have been devised to get around this difficulty by inserting a second mirror to place the focus to one side of the incoming radiation or behind the primary mirror.

10.4.4.1. *The Newtonian Telescope.* The basic system used for most small astronomical reflecting telescopes is a plane mirror located near the focus of the primary mirror (conventionally a paraboloid) and at 45° to the optical axis so that the focus is located immediately outside the incoming beam, as shown in Fig. 10-24.

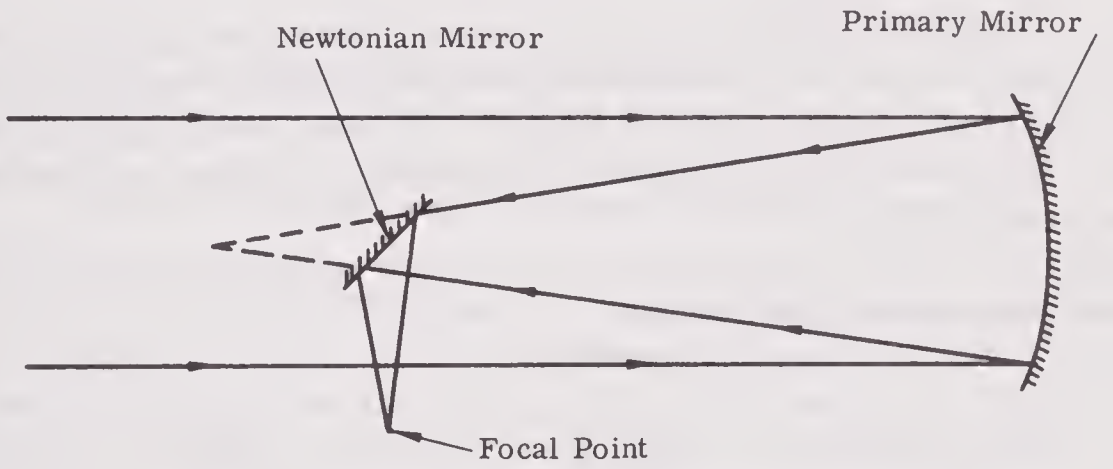


FIG. 10-24. Newtonian telescope.

10.4.4.2. *The Folded Reflector.* The system shown in Fig. 10-25 is the simplest way of locating the focal point behind the primary. A plane mirror, perpendicular to the axis and lying between the primary and its image reverses the direction of the light, so that it passes through the central hole cut in the primary. This places the focal point beyond the mirror where it is more accessible. However, this mirror obscures at least half the aperture.

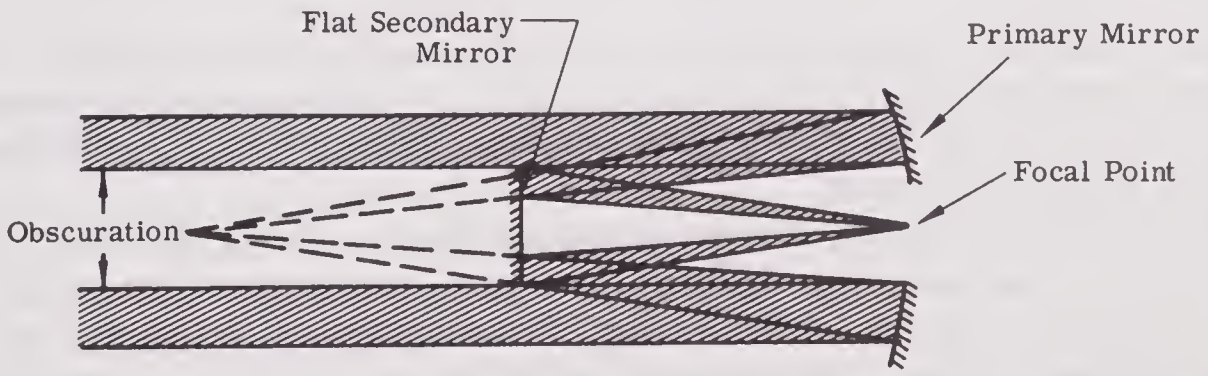


FIG. 10-25. Folded reflector.

10.4.4.3. *The Cassegrainian Objective.* If the secondary mirror of Fig. 10-25 is made convex, the focal length of the objective is increased, and the obscuration of the beam by the secondary mirror is reduced. Figure 10-26 shows a typical Cassegrain.

The original Cassegrainian objective had a paraboloid for the primary and a hyperboloid for the secondary. Both were used in such a way that the image formed by each was free of spherical aberration (see Sec. 10.4.3). The Cassegrain is widely used in infrared work because of the accessibility of its focal point and its relatively low obscuration ratio. By proper selection of "radii," the Petzval surface can be made flat.

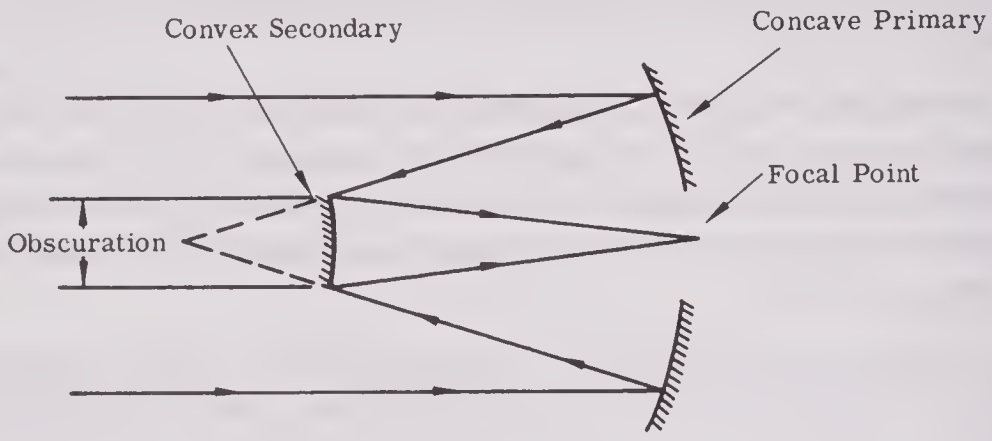


FIG. 10-26. Cassegrainian objective.

10.4.4.4. *The Gregorian Telescope.* The original Gregorian objective consisted of parabolic primary mirror and an elliptical secondary which was placed beyond the focus of the primary mirror so that its "object" and image were located at the foci of the ellipse. Thus, there is a real internal image in the Gregorian, and the final image is erect, as shown in Fig. 10-27.

The Gregorian is rarely used because it is a longer system than the Cassegrainian and offers no compensating advantages.

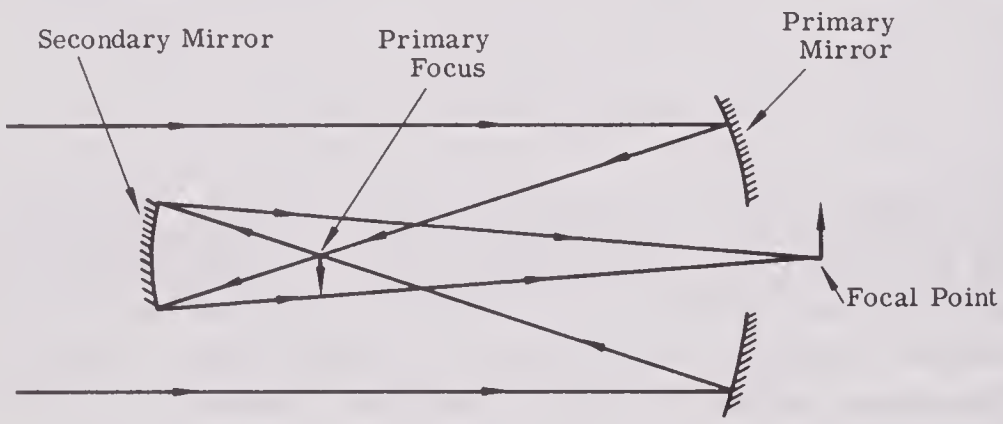
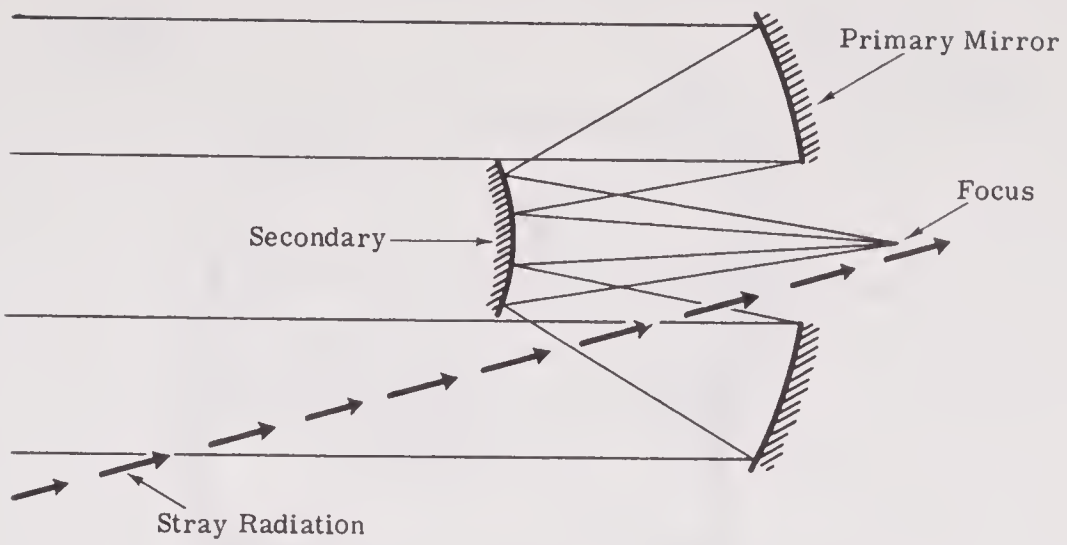


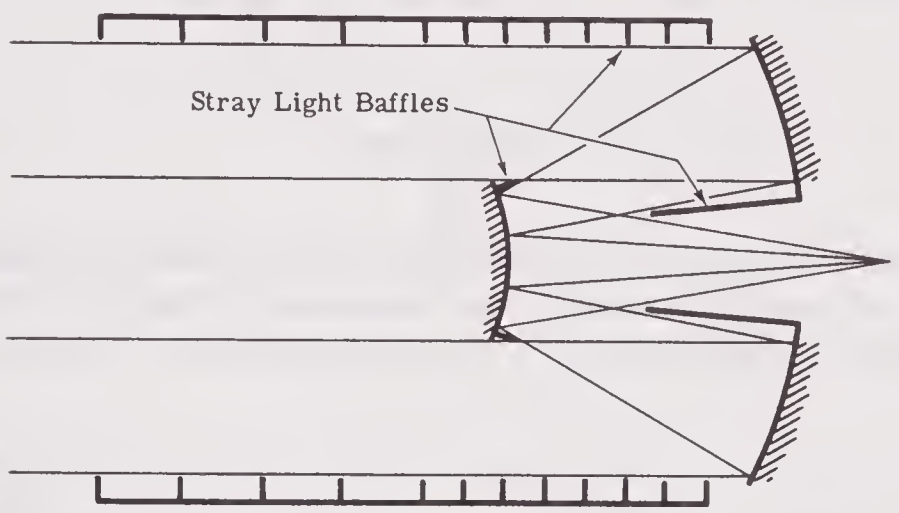
FIG. 10-27. Gregorian telescope.

10.4.4.5. *Baffling of Folded Systems.* All folded systems must be baffled to prevent extraneous radiation from flooding the focal plane. In systems which require a large field coverage and high relative aperture, complete baffling is complex, and in extreme cases may be nearly impossible.

Figure 10-28(*a*) shows a typical Cassegrainian objective. The useful rays are drawn as solid lines. The dashed lines indicate the path of undesired radiation which passes through the system without striking either mirror and which can flood a focal plane detector with stray light. The placement of baffles and sunshades in such a system to cut off stray light is shown in Fig. 10-28(*b*).



(a)



(b)

FIG. 10-28. Folded systems (*a*) before baffling, (*b*) after baffling.

A field lens that is carefully designed to image the secondary mirror (and nothing more) on the detector surface will also eliminate the stray radiation. However, its material may absorb (or reflect) a portion of the desired radiation, and secondary surface reflections of the stray radiation may still get to the detector. Where possible, the field lens plus baffles make a strong combination.

10.5. Catadioptric Objectives

A catadioptric objective combines reflecting and refracting elements. The systems which are most widely used in infrared work are those with relatively thin refracting components, because these absorb the least energy, and, in general, have the least chromatic aberration.

10.5.1. The Schmidt System. The Schmidt objective combines the advantages of the sphere and the paraboloid. As shown in Fig. 10-29, it consists of an aspheric corrector plate at the center of curvature of a spherical mirror. The surface of the corrector is shaped to compensate for the spherical aberration of the reflector. The corrector is located at the center of curvature, as is the aperture stop, so that the system is free of coma and astigmatism as well.

Aspheric surfaces most often chosen are those which have their minimum thickness at the 0.707 or 0.866 zone (of the marginal ray height) depending on whether it is desired to minimize chromatic aberration or optimize the off-axis correction.

The Schmidt system does not achieve perfect imagery because the off-axis ray bundles do not strike the corrector at the same angle as the on-axis bundles. The effect produced

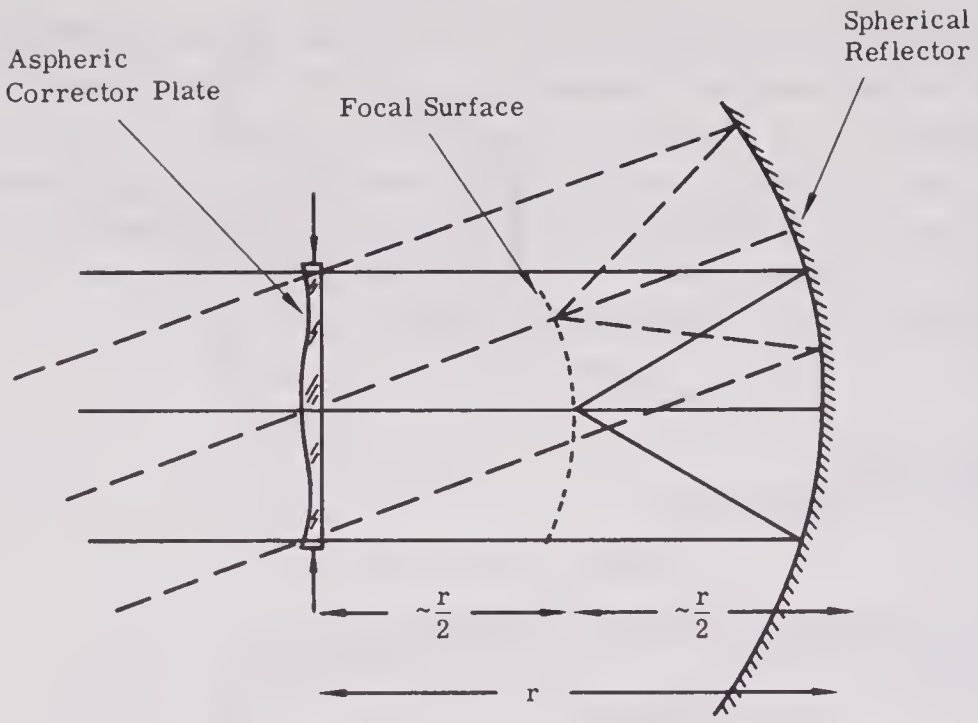


FIG. 10-29. Schmidt system.

is to increase the effect of the corrector, producing an overcorrected condition off axis. (The effect of this on the angular resolution can be estimated from Fig. 10-39 and Eq. 10-51.) The Schmidt performance may be improved somewhat by one or more of the following techniques:

1. Incompletely correcting the axial image to reduce the overcorrection off axis.
2. Using a slightly aspheric primary mirror to reduce the correction load on, and thus the overcorrection introduced by, the corrector.
3. "Bending" the corrector slightly.
4. Using more than one corrector.
5. Using an achromatized corrector.

The aspheric corrector of the Schmidt system is usually easier to fabricate than the aspheric surface of the paraboloid. This is true because the index difference across the corrector, which is usually glass, is only 0.5 compared to the effective index difference of 2.0 across the reflecting surface of the paraboloid, making it only one-fourth as sensitive to fabrication errors. Linfoot [10] and Bouwers [11] have published considerable information about the design and performance of the Schmidt system.

10.5.2. The Mangin Mirror. In the Mangin mirror, Fig. 10-30(*a*), the spherical aberration is corrected by the introduction of a negative lens element in contact with the reflector. For large relative apertures the marginal spherical aberration can be corrected, but a residual remains. A large penalty is paid in the form of the chromatic aberration introduced by the negative refracting element. The Mangin mirror is, however, relatively inexpensive to fabricate and simple to mount.

Figure 10-30(*b*) shows a system with a secondary Mangin mirror, in a sort of Cassegrainian arrangement. This can be used with a Mangin primary to reduce the residual spherical aberration, or with a spherical primary to reduce the cost and weight of the system. The secondary may be designed to be the "power equivalent" of a convex, concave, or plane mirror.

The Mangin mirror may be achromatized by making the negative element an achromatic doublet. Section 10.6 indicates the image quality of the basic Mangin mirror.

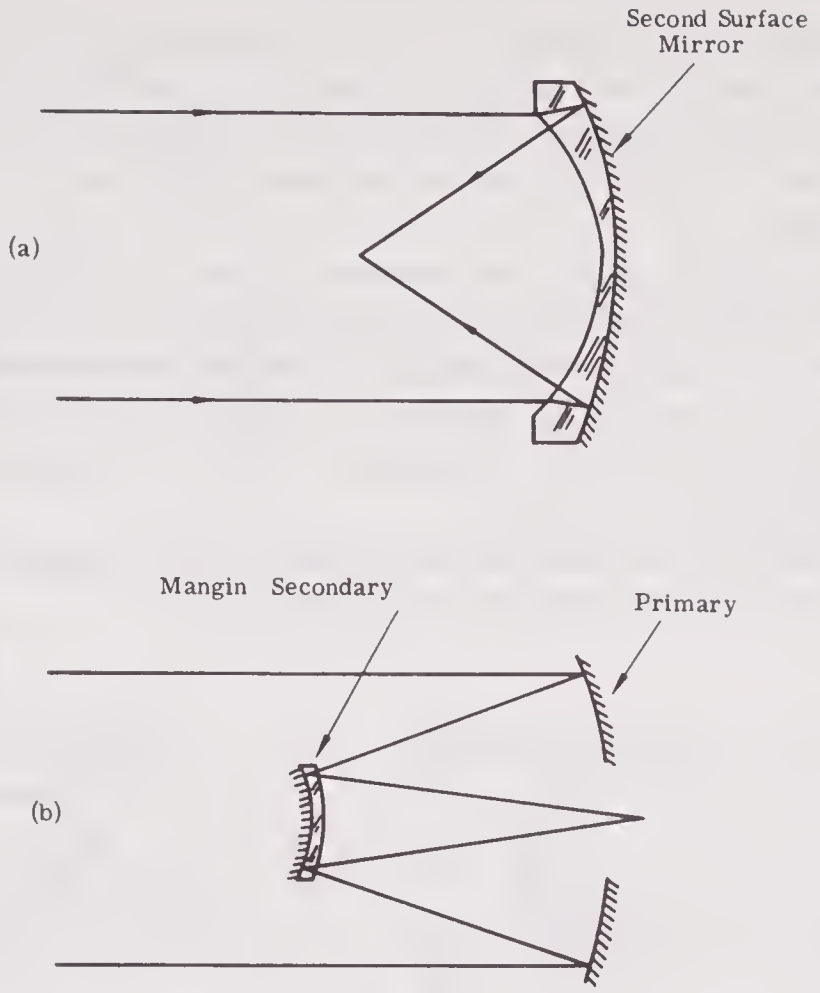


FIG. 10-30. Mangin mirror.

10.5.3. The Bouwers (Maksutov) Objective [11]. The principle of correcting the aberration of a spherical mirror with a negative power-refracting element is further refined in the Bouwers system. By moving the corrector away from the mirror, it is possible to use two degrees of freedom (shape and position) to make a radical improvement over the Mangin mirror.

Figure 10-31 illustrates the concentric system of Bouwers, in which the corrector is placed well away from the mirror and all radii are concentric. This system has the same advantages as the simple sphere with its stop at the center of curvature, in that the image quality is the same for all field angles. The residual spherical aberration of this type of system is low, as is the chromatic aberration, so that an excellent objective results. The design of such a system is easy and fast because there

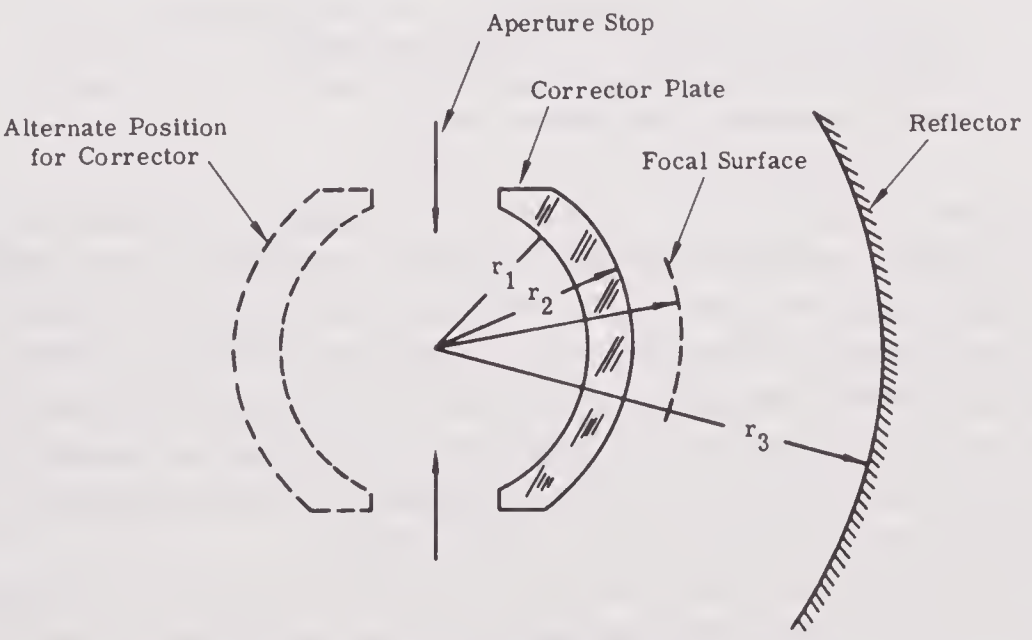


FIG. 10-31. Bouwers concentric system.

are only three variables: the three radii. The technique is to fix r_1 at an arbitrary value; then, for each value chosen for r_2, a value for r_3 can be found for which the marginal spherical aberration is zero. Thus the entire range of designs can be surveyed quickly, because only a few rays need to be traced as all points of the field have the same image quality.

The concentric principle allows the corrector to work identically in either of two positions, behind or before the common center of the system as indicated in Fig. 10-31 by the dashed lines. In its alternate position in front of the common center, the optical effects of the correction are exactly the same; however, this position is used when an extremely high state of correction is desired in order to avoid interference between corrector and focal plane.

The corrector may also be used as the support for a Cassegrainian secondary, as shown in Fig. 10-32 in several different arrangements.

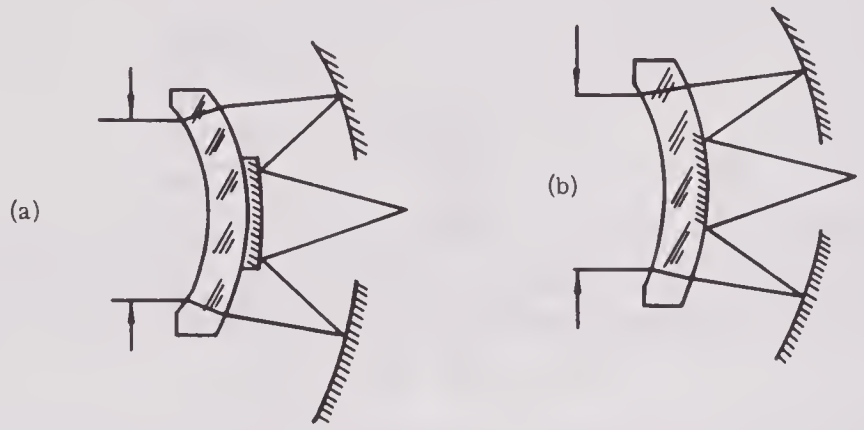


FIG. 10-32. Some arrangements of Bouwers systems as Cassegrainian objectives. An arrangement similar to (*c*) is often used in missile guidance since the corrector makes a reasonably aerodynamic window or dome.

The basic Bouwers system has residuals of chromatic and spherical aberration, which although small are often worth correcting. The chromatic aberration may be reduced by making the corrector a two-element component, with materials and the "buried" surface chosen to achromatize the corrector, as shown in Fig. 10-33. The residual spherical aberration may be eliminated by the use of a Schmidt type aspheric corrector, located at the common center of the system. Since the power of this aspheric corrector is much less than that of a Schmidt, its correction variation with obliquity is correspondingly less, and an excellent system results.

A number of variations on these principles have been produced, using dual correctors (in front and behind center), multiple aspherics, and achromatized aspherics, each of which is designed to improve still further on this excellent system.

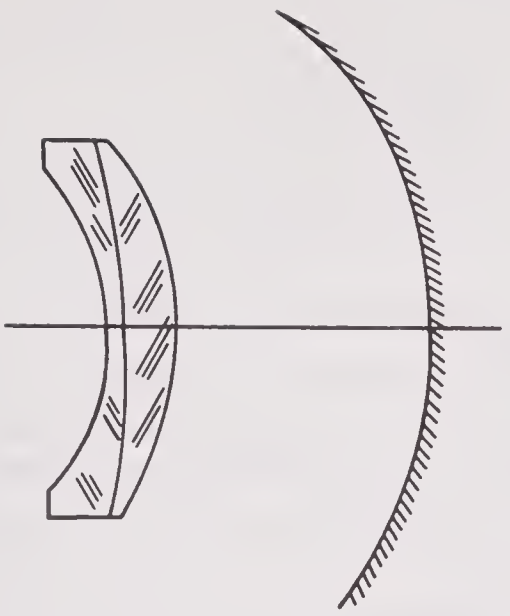


FIG. 10-33. Bouwers system with achromatized corrector plate.

One reason for the popularity of the basic Bouwers system, over and above its excellent performance and easy design, is that, being composed entirely of spherical surfaces, it is relatively inexpensive to make.

The arrangement of the Bouwers with the corrector convex to the incident radiation is widely used, often with a Cassegrainian secondary of one sort or another, for infrared missile guidance systems. The "dome" formed by the corrector then functions as a window for the system as well as a corrector. The corrector need not be concentric when a large well-corrected field is not required. See Fig. 10-34.

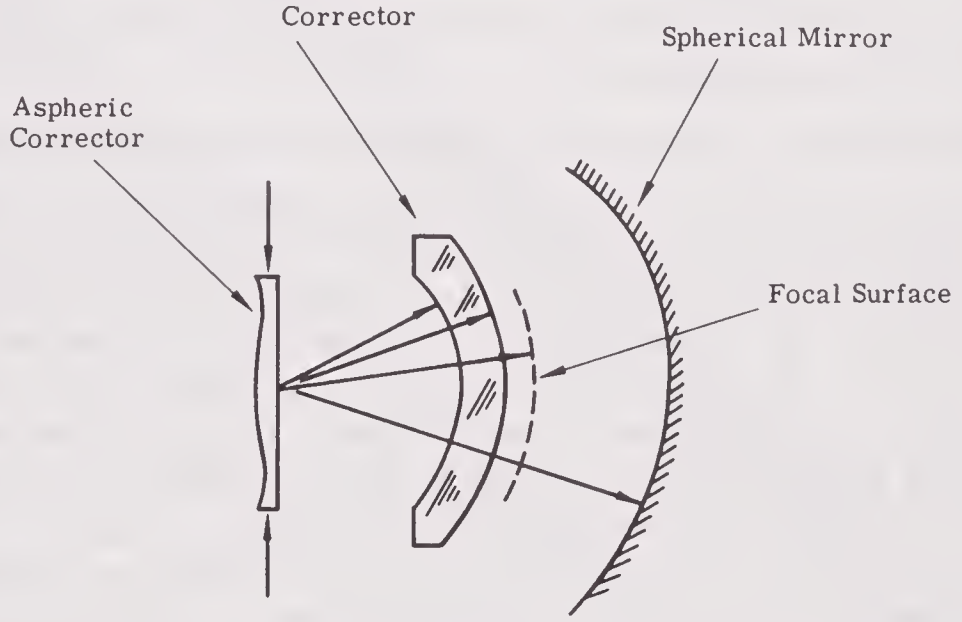


FIG. 10-34. Corrected concentric Bouwers system.

10.5.4. Aberrations of a Plane Parallel Plate in a Convergent Beam. It is often necessary to insert a filter or beam splitter into an optical system in a location at which the beam is not parallel, usually to minimize the size of the insert. This plane parallel plate in a convergent beam introduces aberrations that may significantly affect the performance of the system. The following expressions can be used to evaluate the aberrations introduced by this plate.

Spherical Aberration:

Third Order:

$$SC = \frac{tu^2(n^2 - 1)}{2N^3} \tag{10-40}$$

Trigonometric:

$$LA' = L' - l' = \frac{t}{n}\left(1 - \frac{n\cos u}{\sqrt{n^2 - \sin^2 u}}\right) \tag{10-41}$$

Astigmatism:

Third Order:

$$AC = \frac{tu_p{}^2(n^2 - 1)}{2n^3} \tag{10-42}$$

Coddington:

$$l_s' - l_t' = \frac{t}{\sqrt{n^2 - \sin^2 u_p}}\left[\frac{n^2\cos^2 u_p}{(N^2 - \sin^2 u_p)} - 1\right] = 0.27t \text{ at } u_p = 45° \tag{10-43}$$

Coma-Third Order:

$$CC = \frac{tu^2u_p(n^2 - 1)}{2n^3} \tag{10-44}$$

Longitudinal Chromatic-Third Order:

$$LchC = \frac{t\Delta N}{n^2} = \frac{t(n-1)}{n^2V} \tag{10-45}$$

where t = the thickness of the plate

n = the index of refraction of the plate

u = the convergence angle of the rays to the axis

u_p = the tilt of the plate from normal to the axis

$V = (n-1)/\Delta n$ − Abbé V number.

These equations are derived directly from the surface-contribution equations and the ray-tracing equations of Chapter 9.

10.6. Rapid Estimation of Image Blur Size for Several Optical Systems

The performance required of an optical system usually can be expressed in terms of resolution or image spot size, that is, the size of the smallest blur of light that the system is capable of producing as the image of an infinitesimally small source of light. The minimum size of the image spot is determined either by the size of the system aperture (because of the wave nature of light) or by design characteristics (that is, aberrations) of the optical system. The accompanying figures provide a convenient tool to determine the limiting size of the blur spot for diffraction-limited systems and for several specific optical systems which are widely used.

The figures express the image spot size in one of two ways: as the linear diameter (B) of the image spot, and as the angular diameter (β). The angular diameter β is simply the linear diameter B divided by the effective focal length of the optical system.

10.6.1. Diffraction-Limited Systems. The image size of any optical system is always limited by the wavelength of the radiation involved. In a perfect system (that is, one which has no aberrations or defects of manufacture) this wavelength limitation will determine the minimum size of the blur spot image. The Airy disc, as this blur is called, takes the form of a central blur of light surrounded by alternating light and dark rings of rapidly decreasing intensity. (See also Sec. 9.9.) The central blur contains 84% of the energy, so that we can consider the diameter of the first dark ring about this disc as a conservative value of the effective size of the blur spot. The diameter of this dark ring is given by the expression

$$B = 2.44\lambda(f/\text{no}) \tag{10-46}$$

and the angular size of the blur is given by

$$\beta = 2.44\lambda D^{-1} \text{ rad} \tag{10-47}$$

where λ = the wavelength of the radiation

f/no = the relative aperture or speed of the system (that is, the ratio of focal length to effective diameter)

D = the effective diameter of the system

Figure 10-35 provides a rapid method of evaluating these expressions for the optical systems shown in Fig. 10-36. Paralleling the x and y axes of the chart are a series of spaced lines each labeled with a wavelength from 0.5 to 32 μ. At an angle to the wavelength lines is another set of lines. The x-axis set represents the effective optical diameter in inches; the y-axis set represents the relative aperture. The intersection of an angled line with the appropriate wavelength line indicates the size of the minimum blur for a given system.

Example 1

An optical system with a focal length of 10 in. and a clear aperture of 5 in. working at a wavelength of 2 μ is required. The intersection of the 5-in.-diameter line with the 2-μ wavelength line falls on the abscissa corresponding to an angular blur of

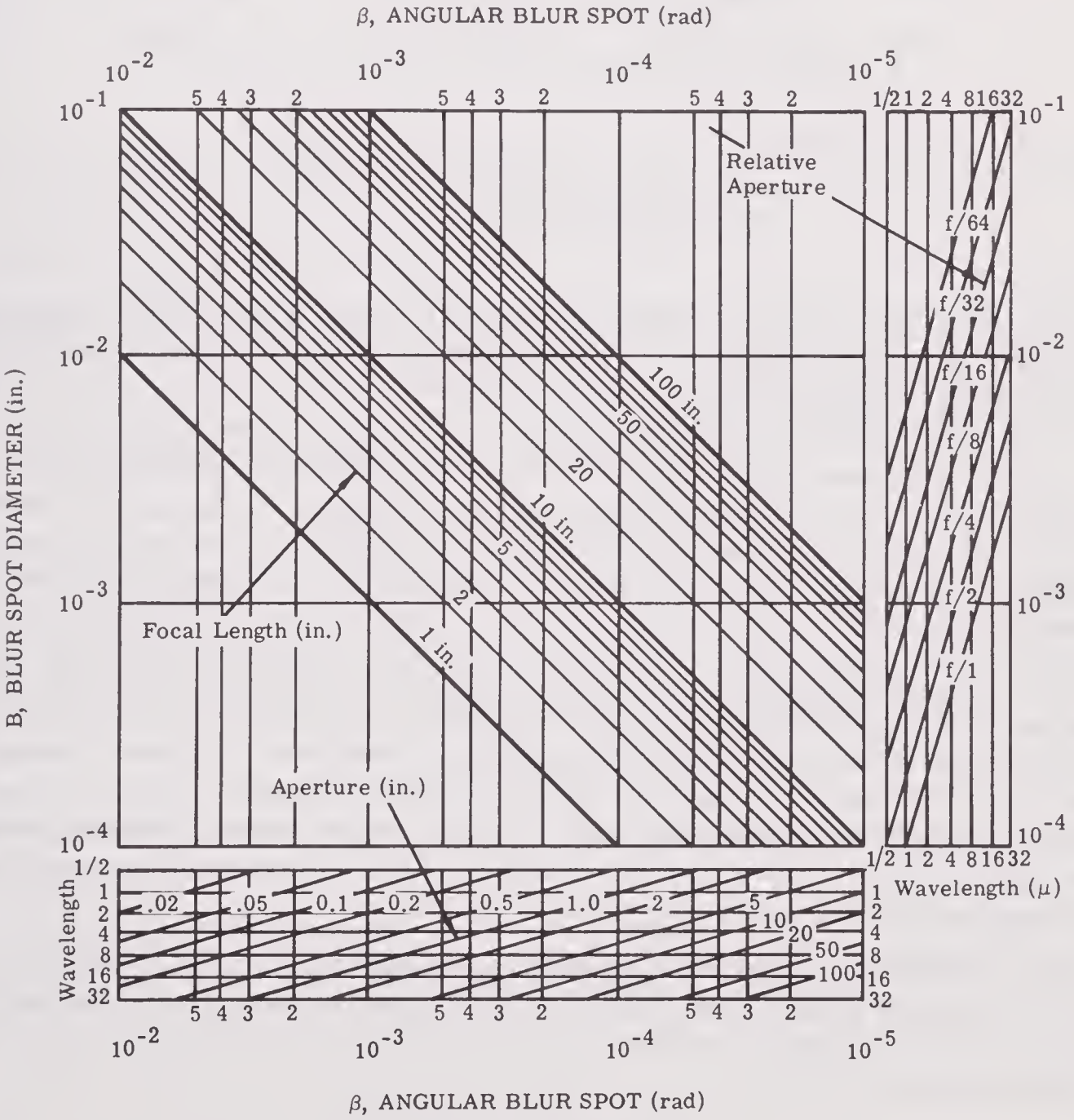
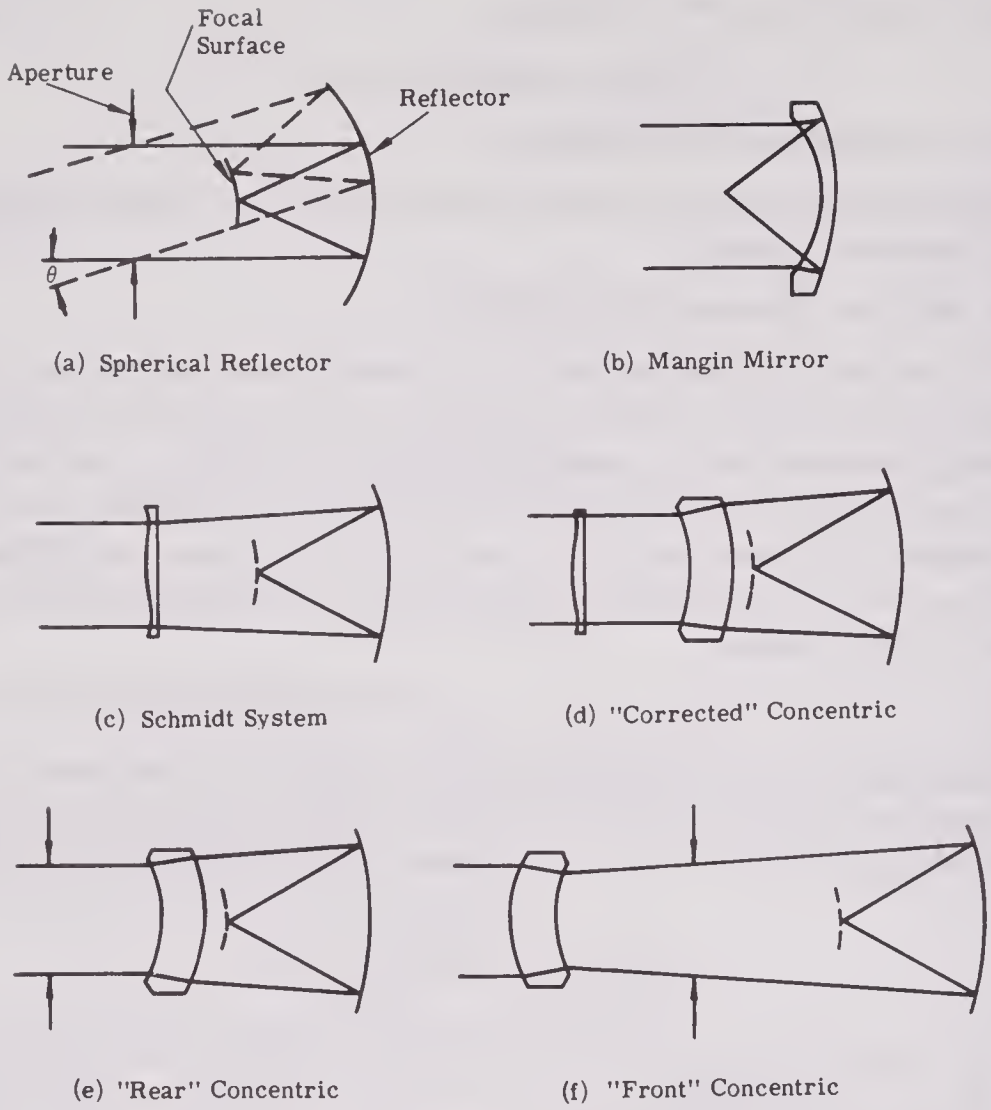


FIG. 10-35. Blur spot size chart for diffraction-limited systems.



(a) Spherical Reflector (b) Mangin Mirror

(c) Schmidt System (d) "Corrected" Concentric

(e) "Rear" Concentric (f) "Front" Concentric

FIG. 10-36. Schematic sections of some common optical systems.

3.8 ± 10^{-5} rad. The speed (relative aperture) of this system is $f/2$ (10-in. focal length divided by the 5-in. diameter). The intersection of the $f/2$ line with the 2-μ line falls on the ordinate corresponding to 3.8×10^{-4} in.

Note that a blur spot diameter of 3.8×10^{-4} in. subtends an angle of 3.8×10^{-5} rad from a distance (focal length) of 10 in. The diagonal focal-length lines of the chart can be used to convert from linear blur diameter to angular blur diameter (and *vice versa*) by following a diameter line across to the intersection with the proper focal length line and then noting the angular size corresponding to the intersection.

Real optical systems usually produce blur spot images which are larger than the diffraction limit described above. Larger blur spots may result from aberrations, that is, the failure of the optical system to produce a perfect focus. In many systems, the size of the aberration blur spot can be expressed as a function of relative aperture (f/no) and field of view (degrees off axis). The following material represents formulas and figures which permit rapid determination of the aberration blur characteristics of several widely used systems.

10.6.2. Spherical Reflector. A simple spherical reflector (Fig. 10-36*a*) with its limiting aperture at the center of curvature has only one image defect, spherical aberration, which is constant over the entire field of view. The minimum angular size of the blur spot is given by

$$\beta = 7.8 \times 10^{-3} \times (f/\text{no})^{-3} \text{ rad} \qquad (10\text{-}48)$$

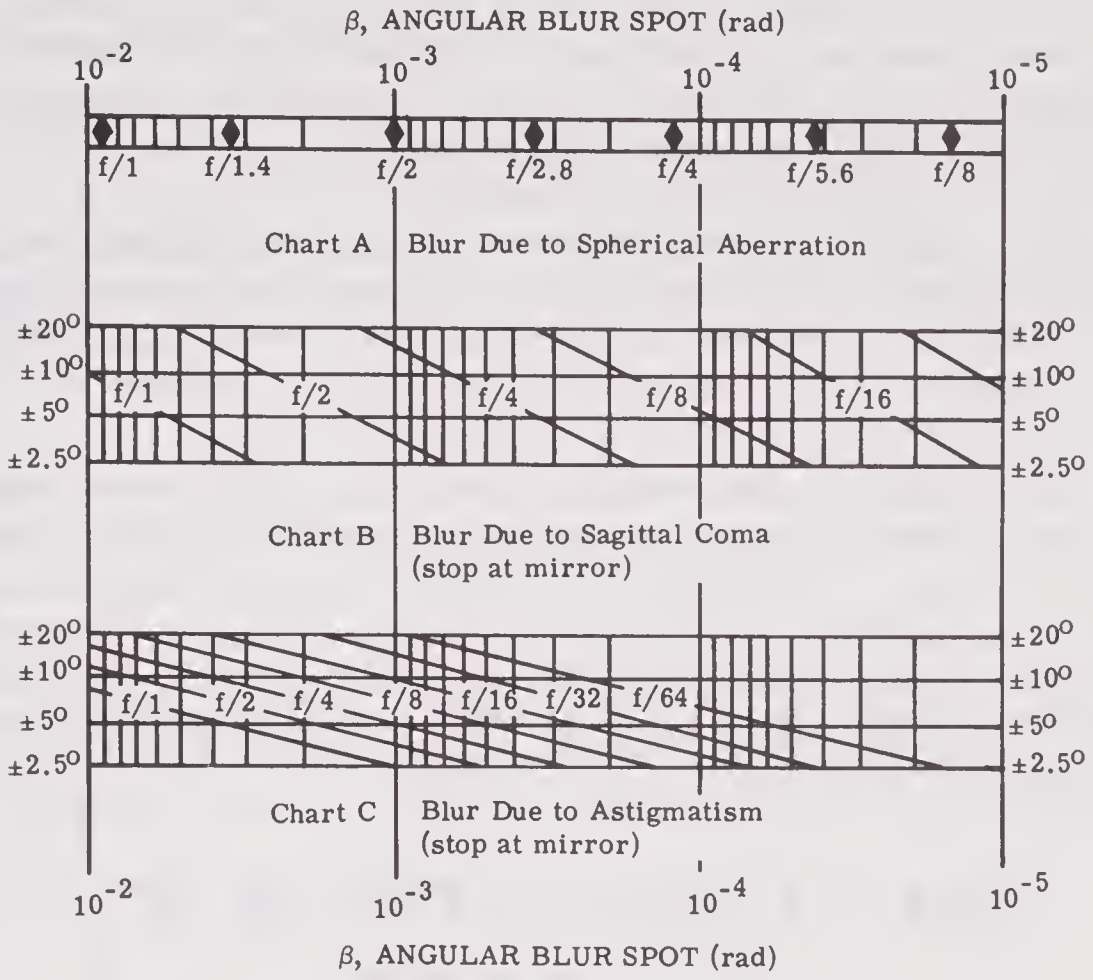


FIG. 10-37. Blur spot size chart for spherical reflector.

for modest apertures (for large apertures the constant is somewhat larger, for example, 9.1×10^{-3} for $f/1$). To evaluate this expression, locate the desired speed (Fig. 10-37(A) and read the corresponding angular blur (β) from the abscissa scale.

When the limiting aperture is not at the center of curvature, coma and astigmatism are present in the off-axis image. In a typical coma blur patch, most of the energy is contained in a small triangular area which is only one-third of the total height of the full coma patch. (See Fig. 9-11 for a sketch of a typical coma patch.) The size of the small triangular (sagittal) coma patch (coma_s) produced by a spherical reflector when the aperture stop is at the surface of the reflector is given by

$$\text{Coma}_s\ \beta = 0.0625 \times \theta(f/\text{no})^{-2}\ \text{rad} \tag{10-49}$$

where θ is the angle (in radians) that the object and image are off axis. To use Fig. 10-37(B) to determine sagittal coma, select the horizontal line representing the angle θ and locate the intersection with the relative aperture line; then read up to find β.

The minimum blur produced by the astigmatism with the stop at the reflector is given by

$$\text{astig. } \beta = 0.5 \times \theta^2 \times (f/\text{no})^{-1}\ \text{rad} \tag{10-50}$$

and Fig. 10-37(C) may be used to evaluate this expression in the same manner as (B).

Example 2

Assume a spherical reflector (stop at the reflector) with a 10-in. focal length, 5-in. clear aperture (relative aperture $f/2$) with a field of $\pm 10°$ (total field 20°). From Fig. 10-37, the angular blur due to

spherical aberration will be 1.0×10^{-3} rad

sagittal coma at 10° will be 2.7×10^{-3} rad

astigmatism at 10° will be 7.6×10^{-3} rad

The interaction of the various aberrations is, in general, difficult to predict precisely. For most purposes, however, it is sufficient to simply add the blurs and accept the sum as a good indication of the final image blur size. Thus, for this example, the indicated angular blur at the edge of the field would be about 11 mrad (11×10^{-3} rad), about ten times as large as that at the center of the field.

If the aperture stop were between the center of curvature and the mirror, the coma and astigmatism blurs could be approximated by interpolation between the zero value obtained with the stop at the center and the values given above. Coma is a linear function of the separation of stop and center of curvature; astigmatism is a function of the square of this separtion.

10.6.3. Mangin Mirror. The Mangin mirror (Fig. 10-36*b*) is a second surface reflector whose refracting first surface is used to correct the spherical aberration of the reflecting surface. The refraction introduces chromatic aberration, however, and limits the improvement attainable by this technique. Figure 10-38 indicates the angular blur of a Mangin mirror due to residual spherical aberration and chromatic aberration. The sagittal coma blur of the Mangin is about half that of a similar spherical mirror.

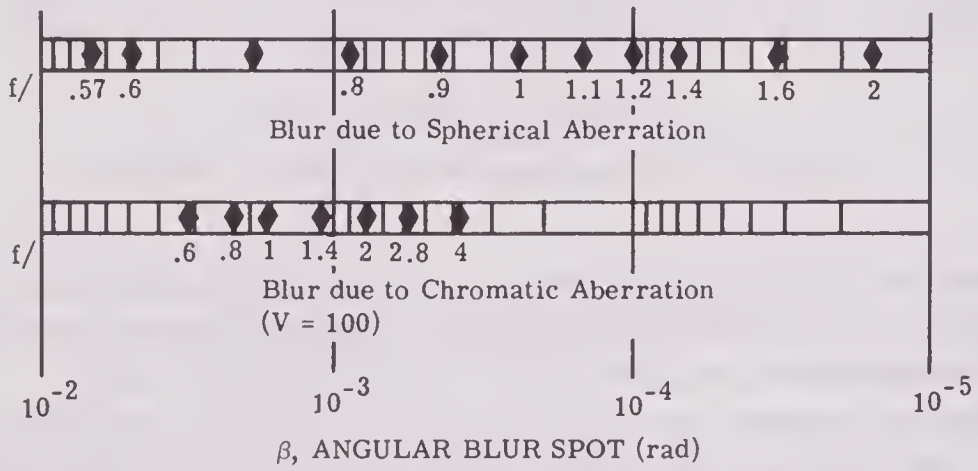


FIG. 10-38. Blur spot size chart for Mangin mirrors.

Example 3

Figure 10-38 indicates that an $f/2$ Mangin mirror has a spherical aberration blur of 1.5×10^{-4} rad and a chromatic blur of about 8×10^{-4} rad, if the material used has a V value of 100 (see Sec. 10.6.7 on refracting systems for definition of V value). If the system operated in the near infrared (1 to 2.3 μ), the equivalent V value of glass for this region would be about 35 and the chromatic blur spot would be 100/35 times as large, that is, about 2.3×10^{-3} rad. The sagittal coma blur would be about 1.4×10^{-3} rad at 10° off axis, half that indicated in Example 2.

10.6.4. Parabolic Reflector. The parabolic reflector, or paraboloid, is widely used because of its perfect image quality on axis, where the angular blur size is theoretically limited only by diffraction as discussed above in Example 1. For off-axis usage, however, coma and astigmatism are present in the same amounts as in a spherical reflector when the stop is in contact with the mirror. (Note that the coma of a parabolic reflector is constant and not a function of stop position, but the astigmatism is zero when the stop is located at the focal plane, that is, one focal length before the mirror.) Thus the angular blur size for a parabolic reflector can be determined from the same equations (10-49 and 10-50), and the same chart (Fig. 10-37) used for a spherical reflector. If we substitute a paraboloid of the same dimensions in Example 2, the blur on axis would be that of Example 1, or 3.8×10^{-5} rad at a wavelength of 2 μ,

and the off-axis blurs caused by coma and astigmatism would be 2.7×10^{-3} and 7.6×10^{-3} rad, respectively, as in Example 2.

10.6.5. Schmidt System. The Schmidt system (Fig. 10-36*c*) combines the perfect axial image quality of the paraboloid with the uniform image quality over a wide field. This is accomplished by the use of a spherical primary mirror with a thin refracting aspheric corrector plate located at the center of curvature of the primary. Thus, the Schmidt system is diffraction limited on axis; the size of the angular blur spot off axis is given by

$$\beta = 0.0417 \times \theta^2 \times (f/\text{no})^{-3} \tag{10-51}$$

A Schmidt system with an $f/2$ relative aperture would have a 1.6×10^{-4}-rad blur at 10° off axis. Figure 10-39 is used in the same manner as Fig. 10-37(*B*) and (*C*).

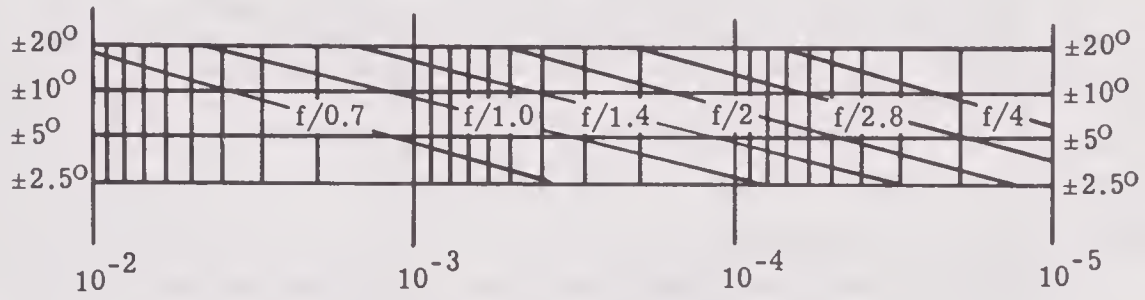


FIG. 10-39. Blur spot size chart for Schmidt systems.

10.6.6. Catadioptric Systems. Systems which combine refracting and reflecting elements are known as catadioptric systems (Fig. 10-36*d*, *e*, and *f*). Perhaps the most useful of these is the concentric Bouwers system. Because of its concentric character the image blur size is the same over the entire field of view. Figure 10-40 presents the limiting characteristics for two forms of the concentric Bouwers and also for the corrected concentric Bouwers, which is a concentric Bouwers with a Schmidt-type corrector at the center of curvature.

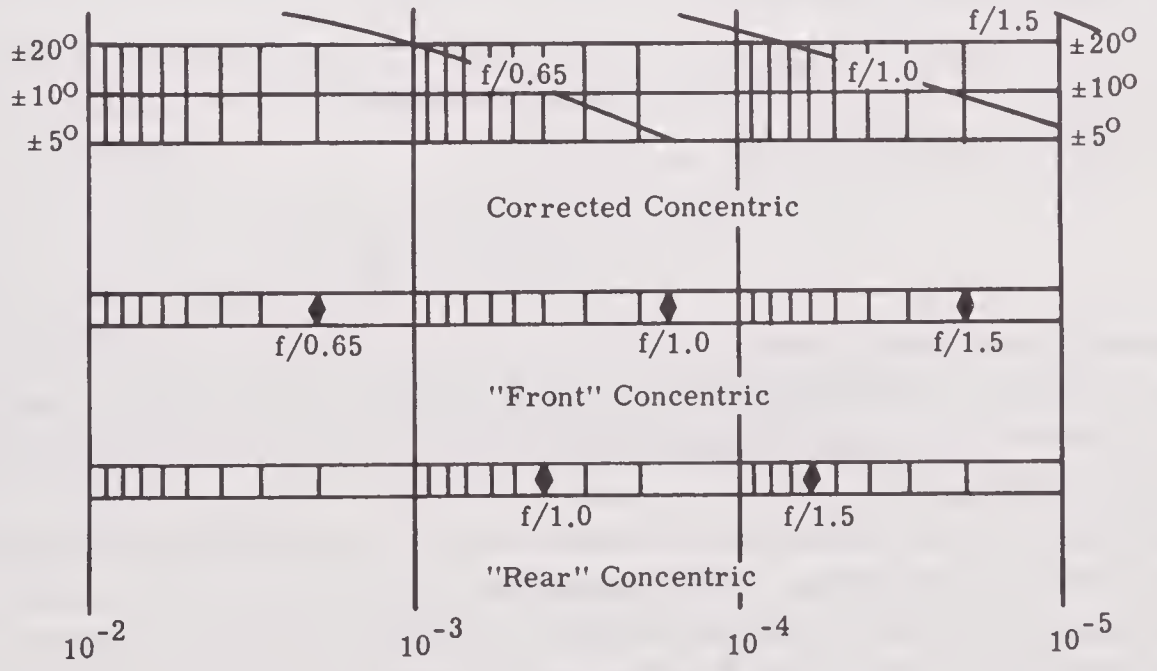


FIG. 10-40. Blur spot size chart for concentric Bouwers systems.

10.6.7. Refracting Systems. Because of the complexity of refracting systems, only the most elementary types lend themselves to anything approaching a complete analysis. The paragraphs below consider only a single thin element with the aperture stop at the element.

The spherical aberration of a thin lens is a function of its index of refraction (a characteristic of its material) and its shape. Figure 10-41(*A*) indicates the angular blur due to spherical aberration of a lens shaped to minimize aberration, when the surfaces of the lens are spherical. (Spherical aberration can be eliminated by the use of aspheric surfaces.) The example of an $f/2$ optical system as a lens would thus have a blur of 8×10^{-3} rad if made of glass ($n = 1.5$) or 1×10^{-3} if made of germanium ($n = 4$).

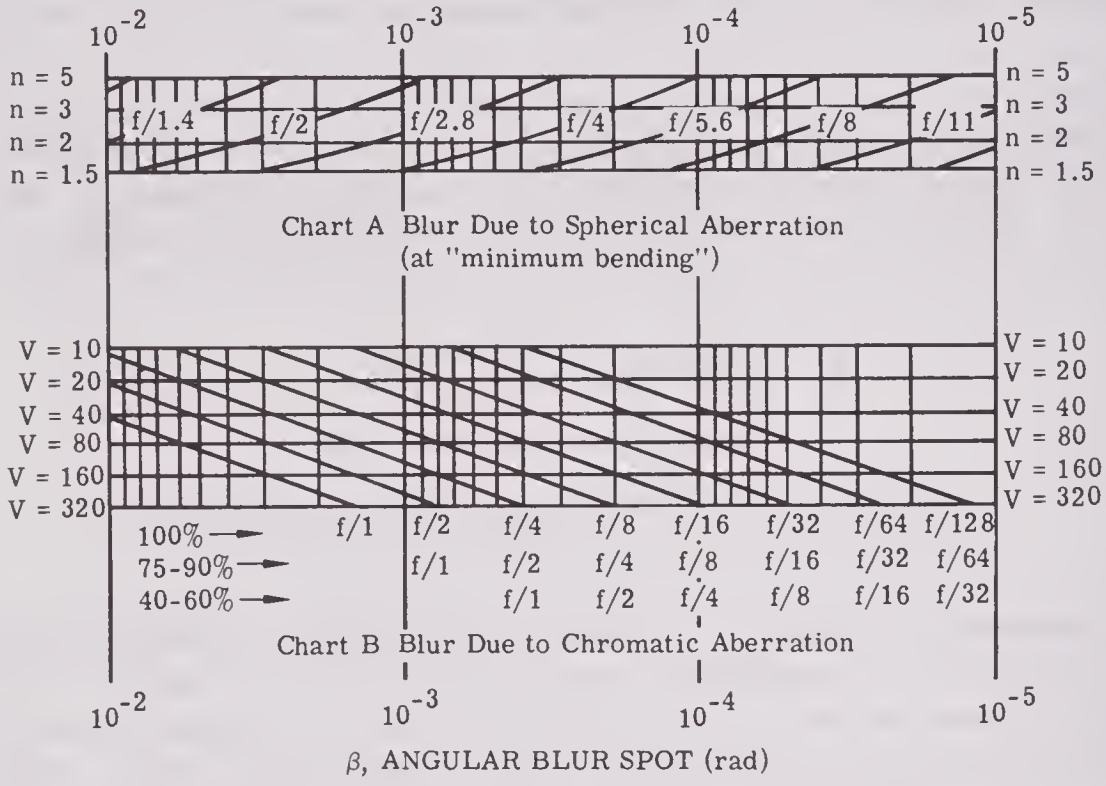


FIG. 10-41. Blur spot size chart for single refracting element.

The coma blur of a thin lens, shaped for minimum spherical aberration, with the stop in contact with the lens, is given by

$$\text{coma}_s\, B = \frac{\theta}{16(n+2)(f/\text{no})^2} \tag{10-52}$$

The coma is zero if the stop position is given by

$$l_{pr} = 2(n-1)^2/\phi n(1-4n) \tag{10-53}$$

The astigmatism of a lens in contact with the aperture is given by Eq. (10-53) so that, for lenses of "minimum bending," the β as given by Fig. 10-37(*C*) can be used as an indication of the off-axis image characteristics, since the coma of a minimum bending lens is usually quite small.

A simple lens also has chromatic aberration, which is the difference in focal position for light of various wavelengths. This results from a variation of the refractive index of the material of the lens with wavelength. A sort of chromatic "figure of merit" for optical materials is the Abbé V value, which when generalized for use at wavelengths other than the visual, is expressed as

$$V = (n_m - 1) \times (n_s - n_l)^{-1} \tag{10-54}$$

where n is the refractive index and the subscripts m, s, and l refer to the index at the middle, short, and long wavelengths of the spectral band of sensitivity of the detecting device used with the optical system. Suppliers' catalogs should be consulted for exact index values, from which V may be computed. For optical glass some typical values are:

Visible spectrum	$V = 30$ to 65
Lead sulfide region 0.4 to 2.5 μ	$V = 9$ to 15
Near infrared 1 to 2.3 μ	$V = 30$ to 40

The angular blur due to chromatic aberration is given by

$$\text{chromatic } \beta = 0.5V^{-1}(f/\text{no})^{-1} \tag{10-55}$$

and can be found by use of Fig. 10-37(B) when V and f/no are known. Note that the first row of f/no values indicates the diameter of the blur containing 100% of the energy. The second row indicates the blur containing 75 to 90% of the energy, and the third row the blur containing 40 to 60% of the total energy. The ordinate corresponding to the intersection of the V value lines and the appropriate f/no line indicates the value of the angular blur spot β.

10.6.8. Summary of Equations.

Diffraction limit

$$\beta = 2.44 \times \lambda \times D^{-1} \tag{10-56}$$

Spherical reflector

Spherical aberration

$$\beta = 0.0078\ (f/\text{no})^{-3} \tag{10-57}$$

Sagittal coma

$$\beta = 0.0625 \times \theta \times (f/\text{no})^{-2} \tag{10-58}$$

Astigmatism

$$\beta = 0.5 \times \theta^2 \times (f/\text{no})^{-1} \tag{10-59}$$

Coma and astigmatism are zero when stop is at the center of curvature. Coma is a linear and astigmatism is a square function of stop to center-of-curvature distance.

Parabolic reflector

Astigmatism and coma blurs per Eq. (10-58) and (10-59).

Coma does not vary with stop position.

Astigmatism is zero for stop at focal plane.

Mangin Mirror

Figure 10-38 gives computed spherical and chromatic blurs.

Coma blur is approximately one-half that of Eq. (10-58).

Schmidt System

Total meridional blur

$$\beta = 0.0417\theta^2 \times (f/\text{no})^{-3} \tag{10-60}$$

Bouwers systems

Figure 10-38 gives computed blurs.

Single-element refracting (spherical surfaces)

Spherical

$$\beta = K(f/\text{no})^{-3} \tag{10-61}$$

$$K - f(n) = 0.067 \text{ for } n = 1.5$$
$$0.027 \text{ for } n = 2$$
$$0.0129 \text{ for } n = 3$$
$$0.0087 \text{ for } n = 4$$

Coma

$$\beta = \theta/[16(n+2)\ (f/\text{no})^2] \tag{10-62}$$

Chromatic

$$\beta = 0.5V^{-1}\ (f/\text{no})^{-1} \tag{10-63}$$

Astigmatism per Eq. (10-59).

References

1. A. C. Hardy and F. H. Perrin, *Principles of Optics* (McGraw-Hill, New York (1932)).
2. P. Drude, *The Theory of Optics* Dover, *Publication, Inc.,* New York (1959).
3. M. Born and E. Wolf, *Principles of Optics,* Pergamon, New York (1959) First ed., 148-149.
4. R. C. Jones, *Proc. IRIS,* **6,** 4, 1 (1961).
5. W. L. Wolfe and J. Duncan, *Proc. IRIS,* **6,** 2, 25-28 (1961).
6. R. DeWaard and E. Wormser, *Thermistor Infrared Detectors, Part I, Properties and Developments,* NAVORD 5495, Barnes Engineering Co., Stamford, Conn. (1958).
7. A. E. Murray, *Proc. IRIS,* **6,** 1, 145 (1961).
8. D. E. Williamson, *J. Opt. Soc. Am.,* **42,** 712 (1952).
9. *Optical Light Pipe Design Study,* Final Report, Contract No. HAC P. O. 4-541998-FF-90-3, Santa Barbara Research Center, Goleta, Calif. (1960).
10. E. H. Linfoot, *Recent Advances in Optics,* Oxford University Press, Amen House, London, E. C. 4 (1955).
11. A. Bouwers, *Achievements in Optics* (Elsevier, Netherlands (1950)).

Bibliography

Design Examples of Cassegrain and Schmidt Type Optics, **5,** Institute of Optics, The University of Rochester, Rochester, New York (1963).

Geometrical Optics, **I,** Institute of Optics, The University of Rochester, Rochester, New York (1963).

Image Evaluation Techniques, **3,** Institute of Optics, The University of Rochester, Rochester, New York (1963).

Optical Design Techniques, **2,** Institute of Optics, The University of Rochester, Rochester, New York (1963).

Chapter 11

DETECTORS

T. Limperis

The University of Michigan

CONTENTS

11.1. Introduction 458
11.1.1. Responsive Elements 458
11.1.2. Windows 459
11.1.3. Apertures 459
11.1.4. Dewar Flask 459
11.2. Detector Parameters 459
11.3. Data Enumeration 460
11.4. Test Procedures 501
11.4.1. Determination of NEP 501
11.4.2. Time Constant 504
11.4.3. Frequency Response 504
11.4.4. Pulse Response 505
11.4.5. Spectral Response 508
11.4.6. Noise Spectrum 510
11.4.7. Sensitivity Contours 510
11.4.8. General Comments 512
11.5. Theoretical Limit of Detectivity 512
11.5.1. Derivation of D^* for Photon Noise Limitation 513
11.5.2. System Design Considerations 515

11. Detectors

11.1. Introduction

A detector is a device providing an electrical output that is a useful measure of incident radiation. It includes not only the responsive element but also the windows, limiting aperture and dewar flask.

11.1.1. Responsive Elements. The responsive element is a radiation transducer. It changes the incoming radiant power into an electrical power which can be amplified by the accompanying electronics. The transduction methods can be separated into two groups. In the first group (thermal detectors), the responsive element is sensitive to changes in its temperature, brought about by fluctuations of the incident radiation power. The second group (photodetectors) contains responsive elements which are sensitive to fluctuations in the number of incident photons. The detectors described in this chapter are listed below. A short description of each detection process follows the listing.

Group 1 Thermal Detectors

Thermistors (bolometric)
Thermocouples (thermovoltaic)
Thermopiles (thermovoltaic)
Golay cells (thermopneumatic)

Group 2 Photodetectors

GaAs
$Cu\text{-}Cu_2O$
InAs
InSb
} Photovoltaic

Si
PbS
PbSe
PbTe
Te
InSb
Ge:Au
Ge:Cu
Ge:Hg
Ge:Cd
Ge:Zn
Ge-Si:Zn
Ge-Si:Au
} Photoconductive

InSb
HgTe (5% ZnTe, 5% CdTe)
} Photoelectromagnetic

Various mixtures of alkali earth oxides } Photoemissive

Bolometric: Changes in temperature of the responsive element induced by the incident radiation causes a change in the electrical conductivity of the element. This change in conductivity is monitored electrically.

Thermovoltaic: When the temperature of a junction of dissimilar metals fluctuates because of changes in the level of incident radiation, the voltage generated by the junction will fluctuate.

Thermopneumatic: When the radiation incident on a gas in a chamber increases the temperature (and therefore the pressure) of the gas, the chamber expands, moving a mirror attached to an external wall. This movement is detected optically.

Photovoltaic: A change in the number of photons incident on a *p-n* junction causes fluctuations in the voltage generated by the junction.

Photoconductive: A change in the number of incident photons causes a fluctuation in the number of free charge carriers in the semiconductive material. The electrical conductivity of the responsive element is inversely proportional to the photon number. This change in conductivity is monitored electrically.

Photoelectromagnetic: Photons absorbed at the surface generate charge carriers which diffuse into the bulk and are separated en route by a magnetic field. This separation of charge produces an output voltage which fluctuates according to fluctuations in the number of incident photons.

Photoemissive: Incident photons impart sufficient energy to electrons on the photoemissive surface to liberate them from the material. The freed electrons may be swept to an anode, and the resulting current can be monitored to determine fluctuations in the number of incident photons.

11.1.2. Windows. Window materials are primarily used to restrict the spectral bandwidth to which the detector is sensitive and to form part of the dewar flask. In selecting the window one must consider a variety of material properties. Chapter 8 describes in detail the various window materials and their properties.

11.1.3. Apertures. Apertures are often used to limit the field of view of the detector in order to reduce photon noise. Section 11.5 of this chapter discusses this subject in some detail.

11.1.4. Dewar Flask. Dewar flasks are used to house the coolant, which reduces the detector temperature in order to improve detectivity. The various dewars and associated cryogenic systems are described in Chapter 12.

11.2. Detector Parameters

The various symbols and units used today for describing the performance of infrared detectors are presented in Table 11-1.* Basic noise mechanisms which limit detecting ability are listed in Table 11-2 along with a description of the physical process and the appropriate mathematical expression. For more detailed information or background material on the individual items in these tables, several textbooks [1,2,3] are available which treat the subject more thoroughly and also provide further references to the archival literature. In addition, two detector state-of-the-art studies [4,5] have been published as government publications.

---

*Table 11-1 is a compilation of the symbols and units arrived at over the years by workers in the infrared community. Especially helpful in preparing the table were Prof. H. Levinstein, Dr. R. C. Jones, and W. L. Eisenman.

11.3. Data Enumeration

This section presents data on commercially available infrared detectors. Of primary concern to the system designer are such parameters as:

1. Detector temperature
2. Normalized detectivity
3. Time constant
4. Resistance or impedance
5. Spectral response
6. Noise spectrum
7. Responsive area
8. Responsivity
9. Viewing solid angle
10. Background temperature

Data of this kind are available from the manufacturers' brochures or from either of the two cell-testing facilities: Naval Ordnance Laboratory, Corona, California (NOLC), and Syracuse (New York) University (SU). NOLC and SU have each published a series of reports describing the results of their measurements: *Properties of Photodetectors* [6] and Interim Report on Infrared Detectors [7], respectively. Compilations of data from these series and other sources for each commercially available detector are presented in the data sheets at the end of this section.*

In addition to the parameters listed above, the design engineer must be concerned with the effects of environmental factors on detector performance. Data of this kind are not available from brochures or from the archival literature; one must rely on confirmation from the contractor. The Detector Evaluation and Information Committee of IRIS (Infrared Information Symposia) has compiled minimum environmental specifications (see next page). Most detectors meet these specifications.

---

*Three detectors which are not commercially available but are of interest have been omitted. They include superconducting bolometers, carbon bolometers, and Tl_2S. At least one other, the germanium bolometer, has not reached the state of maturity such that reliable data are available.

IRIS Environmental Specification #101 [8]

A. Storage Temperature Requirements

1. *Minimum Range:* The range of minimum temperature for detector storage shall be +71° C to −55° C.

2. *Accelerated Life Test:* The detectors shall meet minimum specification requirements within 24 hours after storage for 200 hours at 71° C ± 2° C.

3. *Temperature Cycling Tests* (Thermal Shock): Detectors shall meet performance specifications within 24 hours after being subjected seven times (not necessarily consecutive) to the following cycle: The detector temperature shall be raised from room temperature to 71° C ± 2° C and held there for two hours; lowered within five minutes to −40° C, or below, and held there for two hours; raised within five minutes to 71° C ± 2° C and held for two hours; lowered within five minutes to −40° C or lower, and held there for two hours; and then raised to room temperature, completing the cycle.

B. Humidity Requirements

The detector shall meet the performance specifications within 24 hours after spending ten days in an environment having a relative humidity of (+0%, −5%) and a temperature of 38° C ± 2° C.

C. Vibration Requirements

The detector element plus package shall meet the performance specifications after vibration along all three mutually perpendicular axes from 20 cps to 2000 cps at the rate of five minutes per octave with a peak acceleration of 10 *g*.

D. Shock

The detector element plus package shall meet the performance specification after being subjected to 18 shocks (three in each of six directions) in accordance with Paragraph 4.15.5.1. of Specification MIL-E-5272 [9].

E. Microphonics

Microphonism depends greatly on the method of potting. System requirements are very different. Microphonism specifications must be established for the specific application requirements of the user.

F. Acceleration

Details depend upon system specification requirement. Tests shall be made in accordance with Specification MIL-E-5272, Paragraph 4.16.1, when applicable [9].

G. Vacuum Environment

The detector shall meet minimum specification requirements within 24 hours after storage for 200 hours at a pressure below 10^{-6} torr. This specification does not apply to detector elements permanently mounted in a vacuum dewar.

H. Other

The following earth environments do not apply to normal applications of infrared detectors: salt spray, fungus, rain, sand, dust, immersion test, and explosion-proofing.

TABLE 11-1. DETECTOR PARAMETERS

Parameter (Symbol) [Preferred Units]	*Definition*	*Defining Equation*	*Functional Relationship*
Responsive Area (A) [cm^2]	For responsive elements made of thin films or single crystals used in the photoconductive and photoelectromagnetic mode, the responsive area is the region between the electrodes. For photovoltaic detectors and for detectors using integrating chambers, the responsive area is the effective area, A_e.	$A = \text{length} \times \text{width}$ $A_e = \iint_s \frac{R(x,y)\,dx\,dy}{R_{max}}$ where s = aperture area, R_{max} = maximum value of $R(x,y)$, R = responsivity.	The aperture is in the (x,y) plane.
Impedance (Z) [ohm]	The slope of the voltage-current curve at bias voltage E_1.	$Z = \frac{dE}{dI}\Big]_{E_1}$	Z is a function of the bias voltage, the interelectrode capacitance, and the level of background irradiance.
Resistance ($\bar{R}$) [ohm]	The ratio of the dc voltage across the detector to the dc current through it.	$\bar{R} = E_{dc}/I_{dc}$	$\bar{R}$ is a function of the detector temperature and in some cases of Ω and T_B.
Background Temperature (T_B) [°K]	The effective temperature of all radiation sources viewed by the detector exclusive of the signal source.	–	–

Operating Temperature (T) [°K]	For uncooled detectors the operating temperature is simply the ambient temperature, and for cooled detectors it is the temperature of the coolant or the heat sink.	–	–	
Detector Solid Angle (Ω) [steradian]	The solid angle (field of view) from which the detector receives radiation.	$\Omega = \iint_s \left[\int_0^{\pi/2} \int_0^{2\pi} \frac{\cos\theta \sin\theta \, R(x, y, \phi, \theta)}{A R_{max}(0, 0)} \, d\phi \, d\theta \right] dx \, dy$ where $R_{max}(0,0)$ is the maximum value of $R(x,y,0,0)$. ϕ and θ are spherical coordinates with ϕ being the azimuthal angle. The Z axis is normal to the plane of the responsive element. If the responsivity is not a function of ϕ the element is said to have circular symmetry, and $\Omega = \pi \sin^2 \Theta/2$ where Θ is the total cone angle.	–	
RMS Signal Voltage (or Current) $\begin{pmatrix} V_{s,\,rms} \\ \text{(or } I_{s,\,rms}\text{)} \end{pmatrix}$ $\begin{bmatrix} \text{volt rms} \\ \text{(or amp rms)} \end{bmatrix}$	That component of the electrical output voltage (or current) which is coherent with P_s, the input signal radiation power. P_s can be monochromatic or have a blackbody character.	If the incident radiation power is periodic in time, $P_s(t) = P_0 + P_1 \cos(\omega_1 t + \phi_1) + \cdots$ then $V_s(t) = V_0 + V_1 \cos(\omega_1 t + \theta_1) + \cdots$ and if the dc gain of the associated electronics is zero, $V_{s,\,rms} = (2)^{-1/2} V_1$ assuming that $\Delta f << f$ and that f lies in the region of Δf; *i.e.*, $\Delta f = f_a - f_b$; $f_b < f < f_a$.	The signal voltage is a function of electrical frequency (f). For a single-time-constant detector $V_{s,\,rms} = \frac{V_{s,\,rms}]_{f=0}}{(1+\omega^2\tau^2)^{1/2}}$	

TABLE 11-1. DETECTOR PARAMETERS (*Continued*)

Parameter (Symbol) [Preferred Units]	*Definition*	*Defining Equation*	*Functional Relationship*
RMS Noise Voltage (or Current) ($V_{n,\,rms}$ (or $I_{n,\,rms}$)) [volt rms (or amp rms)]	That component of the electrical output voltage (or current) which is incoherent with the radiation signal power. This value is determined with the signal radiation power removed. The fundamental physical mechanisms which contribute to this voltage are described in Table 11-2.	If the dc gain of the associated electronics is zero, $V_{n,\,rms} = (<V_n^2>_{av})^{1/2}$	$V_{n,\,rms}$ is related to the detector area, Δf, f, and in some cases to Ω and T_B.
Spectral Responsivity (R_λ) [$\frac{\text{volt rms}}{\text{watt rms}}$] [or $\frac{\text{amp rms}}{\text{watt rms}}$]	The ratio between the rms signal voltage (or current) and the rms value of the monochromatic incident signal power, referred to an infinite load impedance and to the terminals of the detector.	$R_\lambda = \frac{V_{s,\,rms}}{P_{s_\lambda,\,rms}}$	Responsivity is a function of λ, f, T, and bias voltage.
Blackbody Responsivity (R_{BB}) [Same as above.]	Same as above except that the incident signal radiation power has a blackbody spectrum.	$R_{BB} = \frac{V_{s,\,rms}}{P_{s_{BB},\,rms}}$	Same as above.

Time Constant (τ) [–]	A measure of the detector's speed of response. The alternative equations for τ (next column) become identical if the noise has a flat power spectrum and if the responsivity varies with frequency according to the relation $$R_\lambda = \frac{R_\lambda]_{f=0}}{(1+\Omega^2\tau^2)^{1/2}}$$	(*a*) $\tau = 1/2\pi f_c$ — where f_c is that chopping frequency at which the responsivity has fallen to 0.707 of its maximum value.
		(*b*) τ_p is the time required for the signal voltage (or current) to rise to 0.63 times its asymptotic value. It is measured by the light-pulse method: exposing the detector to a "square-wave" pulse of radiation.
		(*c*) Responsive time constant $$\tau_r = \frac{R_{max}^2}{4\int_0^\infty [R(f)]^2\,df}$$
		(*d*) Detective time constant $$\tau_d = \frac{R_{max}^2}{4\int_0^\infty [D^*(f)]^2\,df}$$
		(*e*) Empirical responsive time constant $$\tau_{rs} = \frac{1}{2\pi}\left\{\frac{[R(f_1)]^2 - [R(f_2)]^2}{[f_2R(f_2)]^2 - [f_1R(f_1)]^2}\right\}^{1/2}$$ f_1 and f_2 must be specified
		(*f*) Empirical detective time constant $$\tau_{ds} = \frac{1}{2\pi}\left\{\frac{[D^*(f_1)]^2 - [D^*(f_2)]^2}{[f_2D^*(f_2)]^2 - [f_1D^*(f_1)]^2}\right\}^{1/2}$$ f_1 and f_2 must be specified

TABLE 11-1. DETECTOR PARAMETERS (*Continued*)

Parameter (Symbol) [Preferred Units]	*Definition*	*Defining Equation*	*Functional Relationship*
Spectral Noise Equivalent Power (NEP_λ) [watt]	That value of monochromatic incident rms signal power of wavelength λ required to produce an rms signal to rms noise ratio of unity. The chopping frequency, the electrical bandwidth used in the measurement, and the detector area should be specified.	$NEP_\lambda = P_{s_\lambda, rms}; \frac{V_{n, rms}}{V_{s, rms}} = \frac{V_{n, rms}}{R_\lambda}$	Depends upon λ, A, f, Δf, and in some cases Ω and T_B.
Blackbody Noise Equivalent Power (NEP_{BB}) [watt]	That value of incident rms signal power (with a blackbody spectral character) required to produce an rms signal to rms noise ratio of unity. The blackbody temperature must be specified along with the detector area, the electrical bandwidth used in the measurement, and the chopping frequency.	$NEP_{BB} = P_{s_{BB}, rms}; \frac{V_{n, rms}}{V_{s, rms}} = \frac{V_{n, rms}}{R_{BB}}$	Depends upon blackbody temperature, A, f, Δf, and in some cases Ω and T_B.
Spectral Detectivity (D_λ) [$watt^{-1}$]	The reciprocal of spectral noise equivalent power. The chopping frequency, the electrical bandwidth used in the measurement, and the detector sensitive area should be specified.	$D_\lambda = 1/NEP_\lambda$	Depends upon λ, A, f, Δf, and in some cases Ω and T_B.
Blackbody Detectivity (D_{BB}) [$watt^{-1}$]	The reciprocal of the blackbody noise equivalent power. The blackbody temperature should be specified, along with the electrical bandwidth used in the measurement, the detector area, and the chopping frequency.	$D_{BB} = 1/NEP_{BB}$	Depends upon blackbody temperature, A, f, Δf, and in some cases Ω and T_B.

Spectral D-Star ($D^*(\lambda, f_0)$) [cm $(cps)^{1/2}$ w^{-1}]	A normalization of spectral detectivity to take in account the area and electrical bandwidth dependence. The chopping frequency (f_0) used in the measurement is specified by inserting it in the parentheses as indicated in the last column. For detectors limited by the fluctuation in arrival rate of background photons, Ω and T_B must be specified.	$D^*(\lambda, f_0) = \sqrt{A\,\Delta f}\, D_\lambda$	For background-noise-limited detectors, $D^*(\lambda, f_0)$ depends upon Ω and T_B.
Blackbody D-Star ($D^*(T_B, f_0)$) [cm $(cps)^{1/2}$ w^{-1}]	A normalization of blackbody detectivity to take in account the detector area and the electrical bandwidth. The chopping frequency (f_0) and the blackbody temperature are specified in the parentheses as indicated. For detectors that are background noise limited, Ω and T_B must be specified.	$D^*(T_{BB}, f_0) = \sqrt{A\,\Delta f}\, D_{BB}$	For background-noise-limited detectors, $D^*(T_{BB}, f_0)$ depends upon Ω and T_B.
Maximized D-Star ($D^*_{\mu m}(\lambda_p, f_0)$) [cm $(cps)^{1/2}$ w^{-1}]	A quantity obtained when the wavelength is λ_p, and the chopping frequency used yields a maximum rms signal to rms noise ratio.	–	Same as for $D^*(\lambda, f)$.
Spectral D-Double Star ($D^{**}(\lambda, f_0)$) $\left[\text{cm } (cps)^{1/2}\ w^{-1}\ sr^{1/2}\right]$	A normalization of $D^*(\lambda, f_0)$ to account for the detector field of view Ω. It is used only when the detector is radiation noise limited. (*Note*: if $\Omega = \pi$, $D^{**} = D^*$.)	$D^{**}(\lambda, f_0) = (\Omega/\pi)^{1/2} D^*(\lambda, f_0)$	–
Peak Wavelength (λ_p) [micron]	The wavelength at which detectivity is a maximum.	–	Depends upon cell temperature and detector material used.

TABLE 11-1. DETECTOR PARAMETERS (*Continued*)

Parameter (Symbol) [Preferred Units]	*Definition*	*Defining Equation*	*Functional Relationship*
Cutoff Wavelength (λ_c) [micron]	The wavelength at which $D^*(\lambda, f_0)$ has degraded to one-half its peak value.	–	Depends upon cell temperature and detector material used.
Responsive Quantum Efficiency (RQE) [–]	The ratio of the number of countable output events to the number of incident photons.	$RQE = N_0/N_p$	Depends upon bias voltage, time constant, and cell geometry.
Detective Quantum Efficiency (DQE) [–]	The square of the ratio of measured detectivity to the theoretical limit of detectivity.	$DQE = \left[\frac{D(\lambda_0)_{measured}}{D(\lambda_0)_{\substack{theoretical \\ limit}}}\right]^2$	–

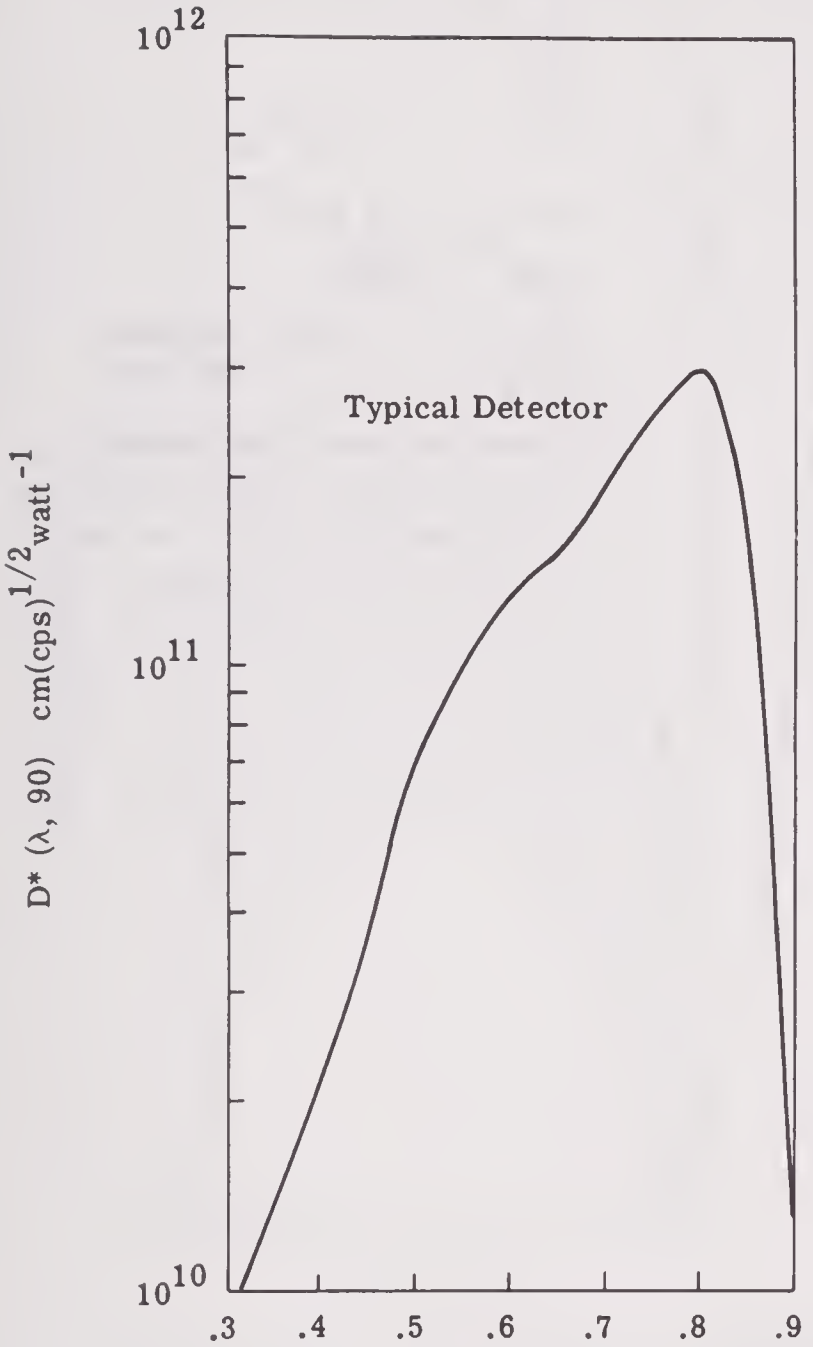
TABLE 11-2. DETECTOR NOISES

Type of Noise	*Physical Mechanism*	*Detectors Concerned*	*Equation for* $V_{n,\,rms}$
Johnson (also called Nyquist or Thermal)	At thermal equilibrium the random motion of charge carriers in a resistive element generates a random electrical voltage across the element. As the temperature of the resistor is increased the mean kinetic energy of the carriers increases, yielding an increased electrical noise voltage.	All detectors.	$V_{n,\,rms} = (4kT\bar{R}\,\Delta f)^{1/2}$ where $k \equiv$ Boltzmann's constant
Temperature	The fluctuations in temperature of the sensitive element, due to either radiative exchange with the background or conductive exchange with the heat sink, produces a fluctuation in signal voltage. For thermal detectors, if the temperature noise is due to radiative exchange with the background, the detector is said to be at its theoretical limit.	All detectors but especially those made of thin films.	For thermal detectors, $\overline{\Delta T^2} = \dfrac{4kT^2K\,\Delta f}{K^2 + 4\pi^2f^2C^2}$ where $\overline{\Delta T^2} \equiv$ mean square temperature fluctuations, $K \equiv$ thermal conductance, $C \equiv$ heat capacity. The relation between $\overline{\Delta T^2}$ and $V_{n,\,rms}$ should be determined for each detector.
Modulation (or $1/f$)	The mechanism is not well understood. As its name implies, it is characterized by a $1/f^n$ noise power spectrum, where n varies from 0.8 to 2.	All detectors but predominately thermal detectors.	$V_{n,\,rms} = \bar{R}I_{dc}C\left(\dfrac{\Delta f}{Ad}\right)^{1/2}\left(\dfrac{1}{f}\right)^n$ where $I_{dc} \equiv$ bias current, $C \equiv$ constant, $d \equiv$ cell thickness

Generation-Recombination (G-R)	Statistical fluctuations in the rate of generation and in the rate of recombination of charged particles in the sensitive element result in an electrical noise. These fluctuations can be caused by charge-carrier-phonon interactions or by the random arrival rate of photons from the background. If the background photons are the prime contributors to the fluctuation in G-R rates, then the noise is often called photon, radiation, or background noise.	All photodetectors.	$V_{n,\,rms} = \bar{R} I_{dc} \left[\frac{2\tau\ \Delta f}{n(1 + 4\pi^2 f^2 \tau^2} \right]^{1/2}$ where n = charge-carrier density. For photovoltaic detectors, the value of $V_{n,\,rms}$ is smaller by a factor of $\sqrt{2}$ since fluctuations occur only in generation, not in recombination.
Shot	Because of the discreteness of electronic charge, the current flowing through the responsive element is the result of current pulses produced by the individual electrons and/or holes.	Photovoltaic detectors and thin-film detectors.	$V_{n,\,rms} = \bar{R}(2qI_{dc}\,\Delta f)^{1/2}$ where q is the charge of an electron.

GaAs

Detector Temperature 300°K

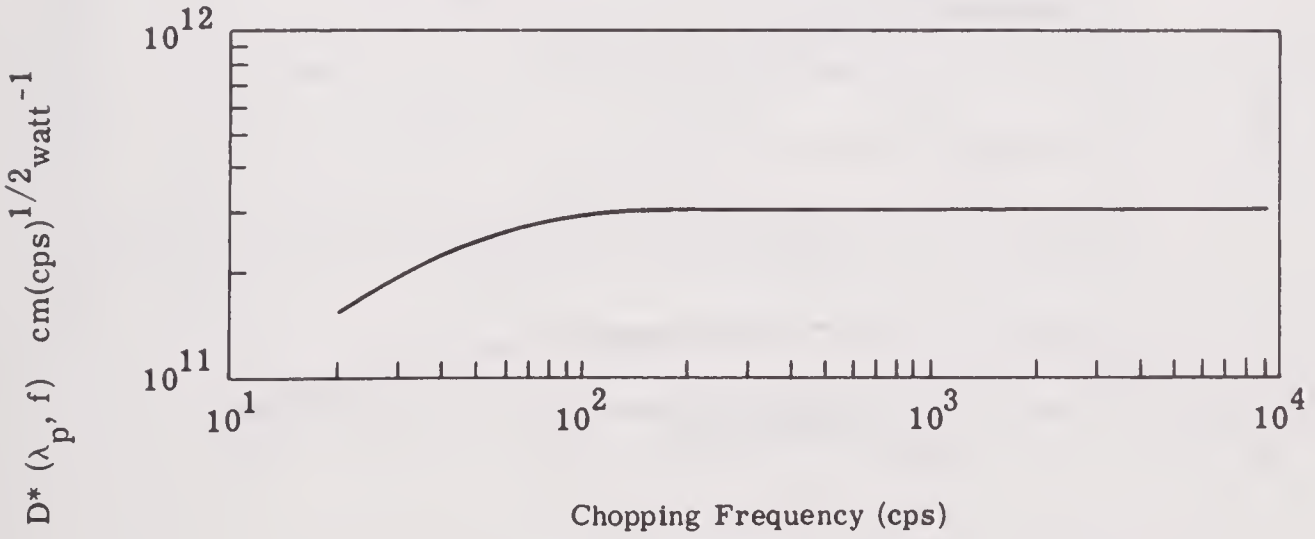


T_B	300°K
τ	<1 μsec
$\bar{R}$	
z	~6×10^6 Ω
A	~10^{-2} cm^2
fov	180°
$R_{\lambda_{max}}$	10^6 volt rms / watt rms

Operating mode: photovoltaic

Limiting noise: 1/f for f <200 cps and shot noise for f >200 cps.

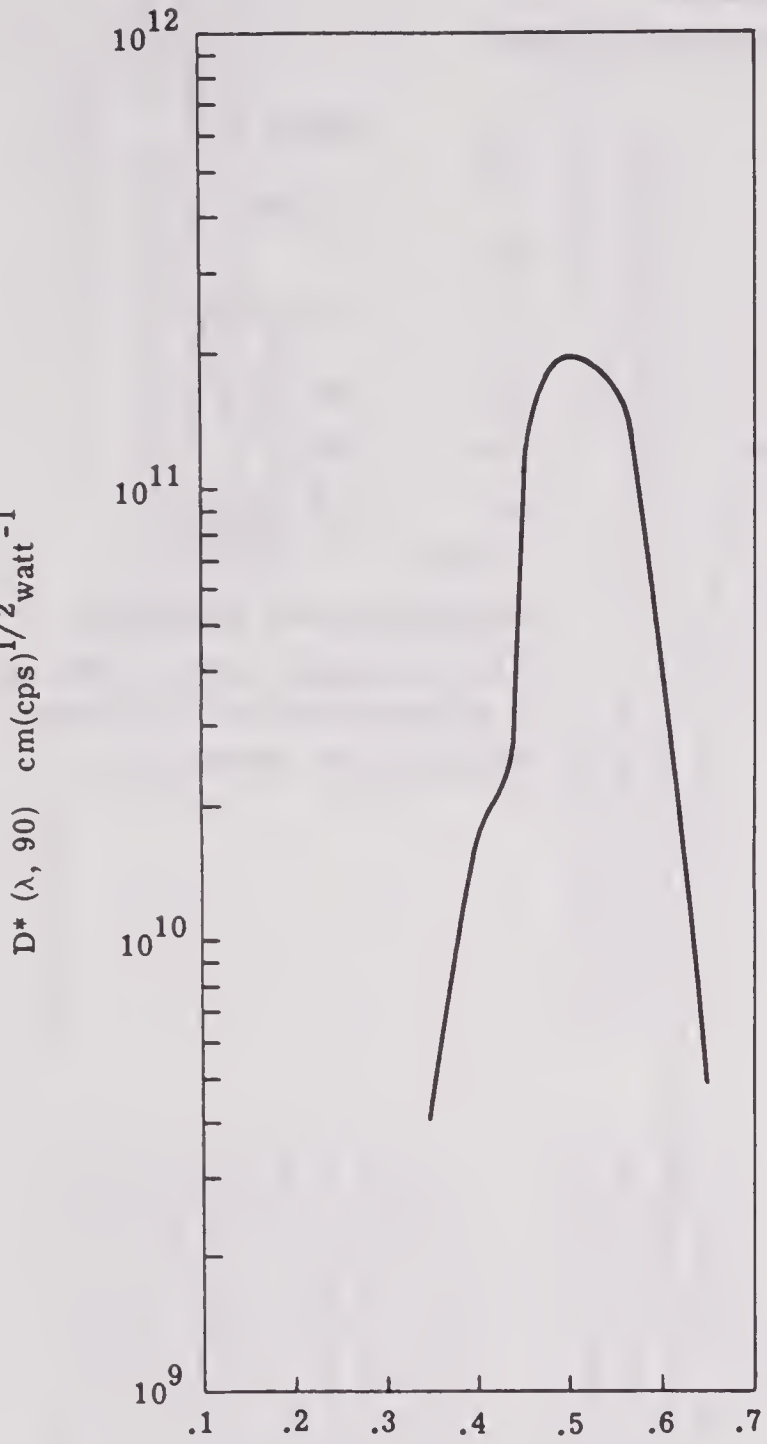
Manufacturer: Philco Corp.



DATA SHEET 1

$Cu\text{-}Cu_2O$

Detector Temperature 300°K

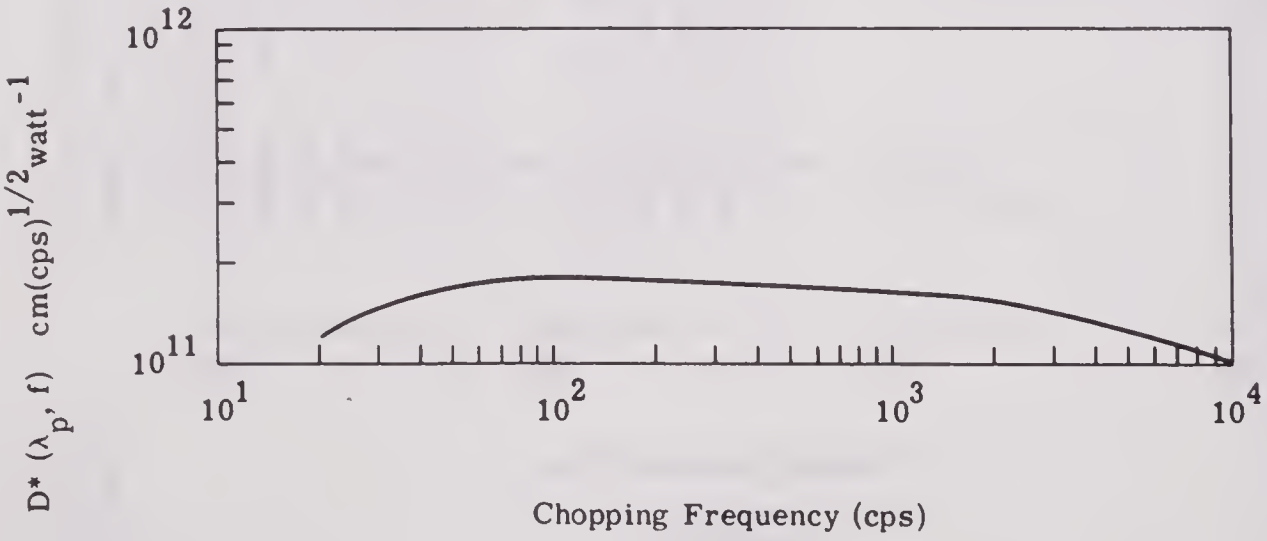


T_B	300°K
τ	
$\overline{R}$	
Z	$\sim 1 \times 10^6\ \Omega$
A	$\sim 4 \times 10^{-2}$
fov	180°
$R_{\lambda_{max}}$	$\sim 5 \times 10^4\ \frac{\text{volt rms}}{\text{watt rms}}$

Operating mode: photovoltaic

Limiting noise: 1/f for f < 100 cps and shot noise for f > 100 cps.

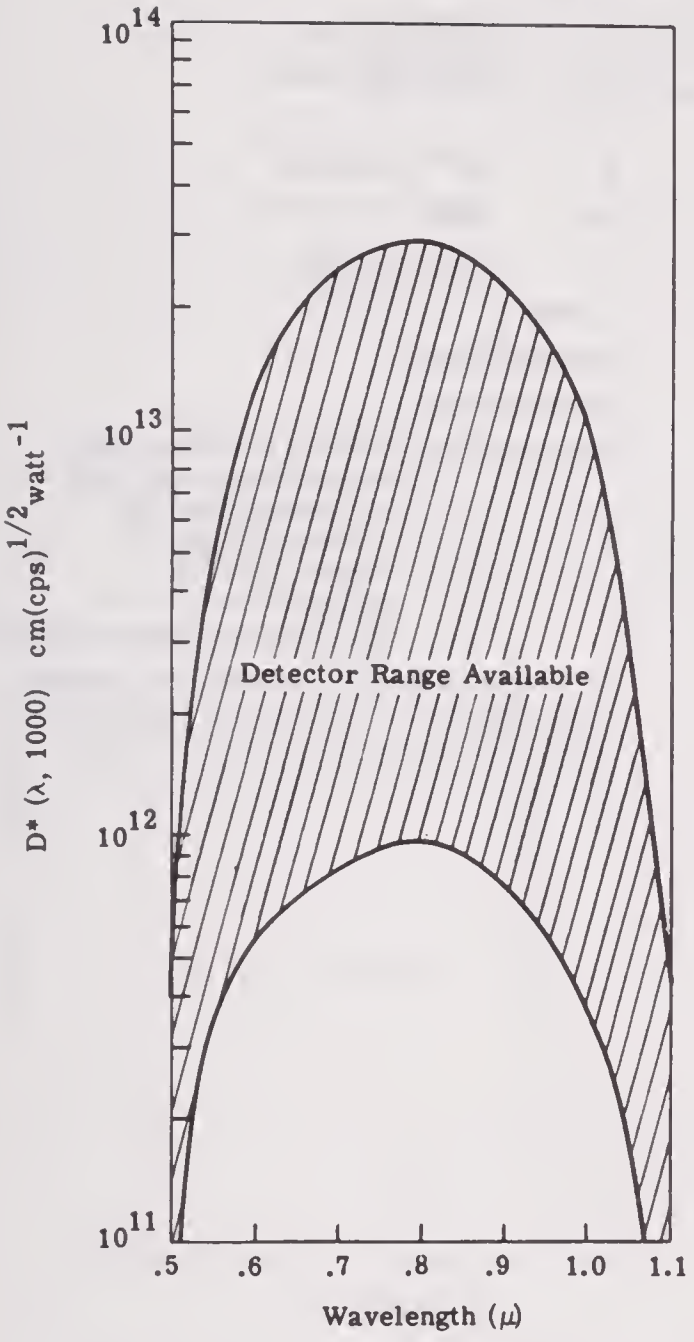
Manufacturer: Philco Corp.



DATA SHEET 2

Si*

Detector Temperature 78°K



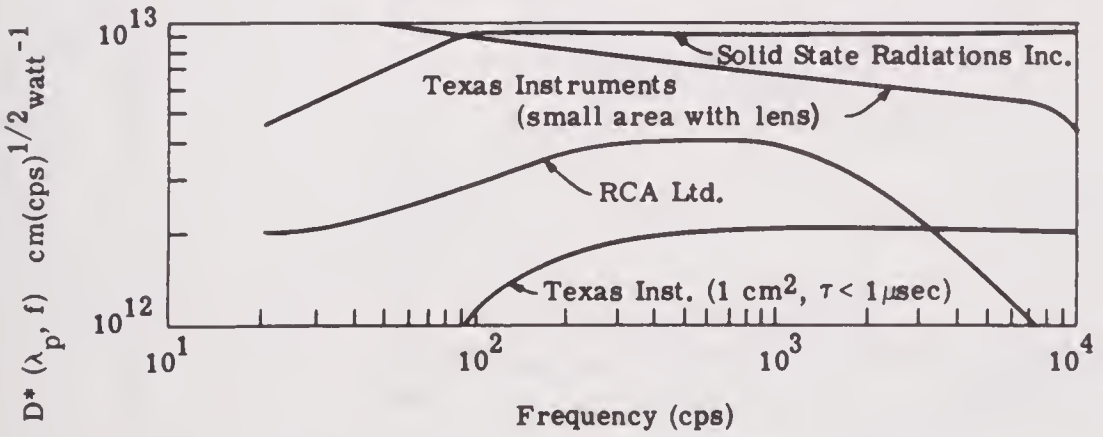
T_B	300°K
τ	<1 to 300 μsec (often limited by the $\bar{R}C$ value of the sensitive element)
$\bar{R}$	~10^8 Ω (for photoconductive detectors)
Z	50 kΩ to 10^7 Ω (for photovoltaic detectors)
A	up to 1 cm^2
fov	180°
$R_{\lambda_{max}}$	10^6 $\frac{\text{volt rms}}{\text{watt rms}}$

Operating mode: photoconductive and photovoltaic

Limiting noise: Johnson noise of the load resistor and shot noise for the photovoltaic mode. G-R and Johnson noise predominate in the photoconductive mode.

Manufacturers: Texas Instruments
Solid State Radiations Inc.
RCA Ltd.
International Rectifier Corp.

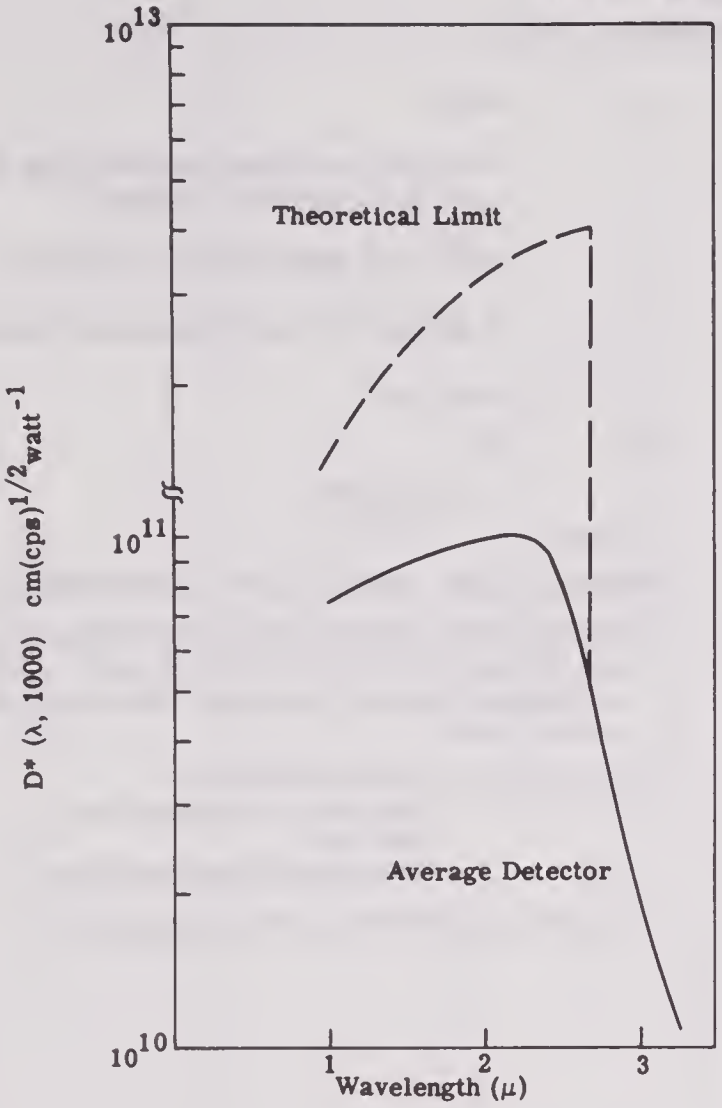
*Lenses are supplied by some manufacturers.



DATA SHEET 3

PbS*

Detector Temperature 300°K



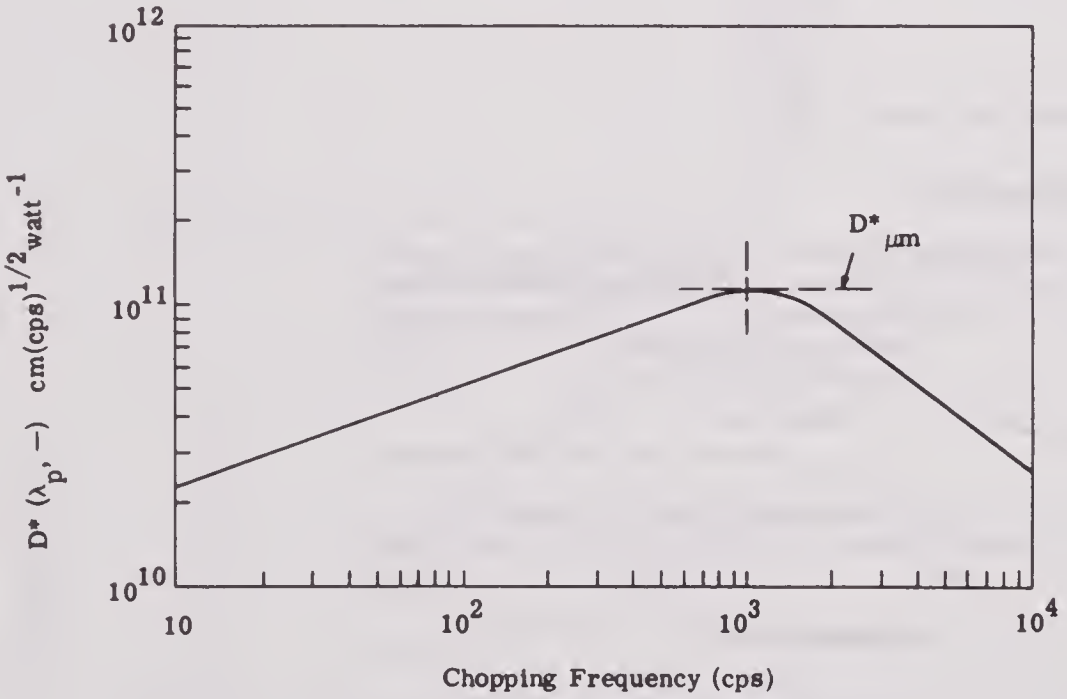
T_B	300°K
τ	1000-5000 μsec
$\overline{R}$	.5-20 MΩ/square
Z	
A	10^{-6} to 10^{0} cm^2
fov	180°
$R_{\lambda_{max}}$	~10^6 $\frac{\text{volt rms}}{\text{watt rms}}$

Operating mode:

Limiting noise:

Manufacturers: Infrared Industries (IRI)
Eastman Kodak Co. (EK)
Electronics Corp. of America (ECA)
Tupper Trent Co.
Mullanrd Ltd.
Santa Barbara Research Center

Tradeoffs between parameters are available upon request from the manufacturer (e.g., τ may be decreased at the expense of D).



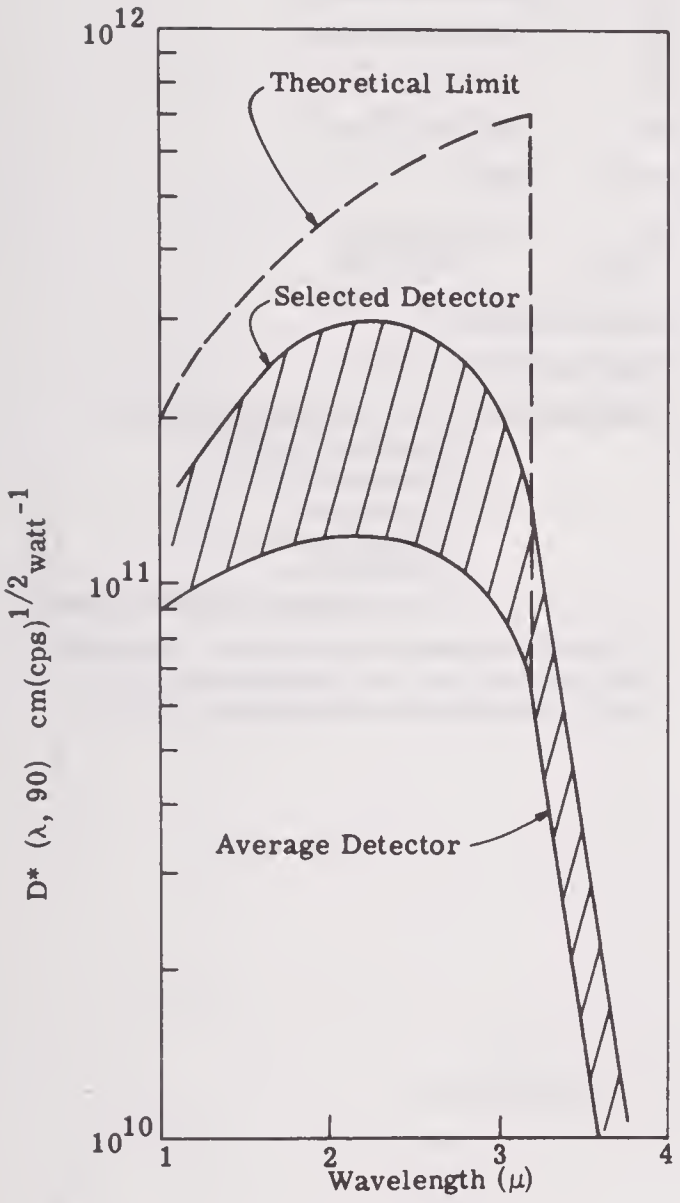
DATA SHEET 4

PbS*

Detector Temperature 195°K

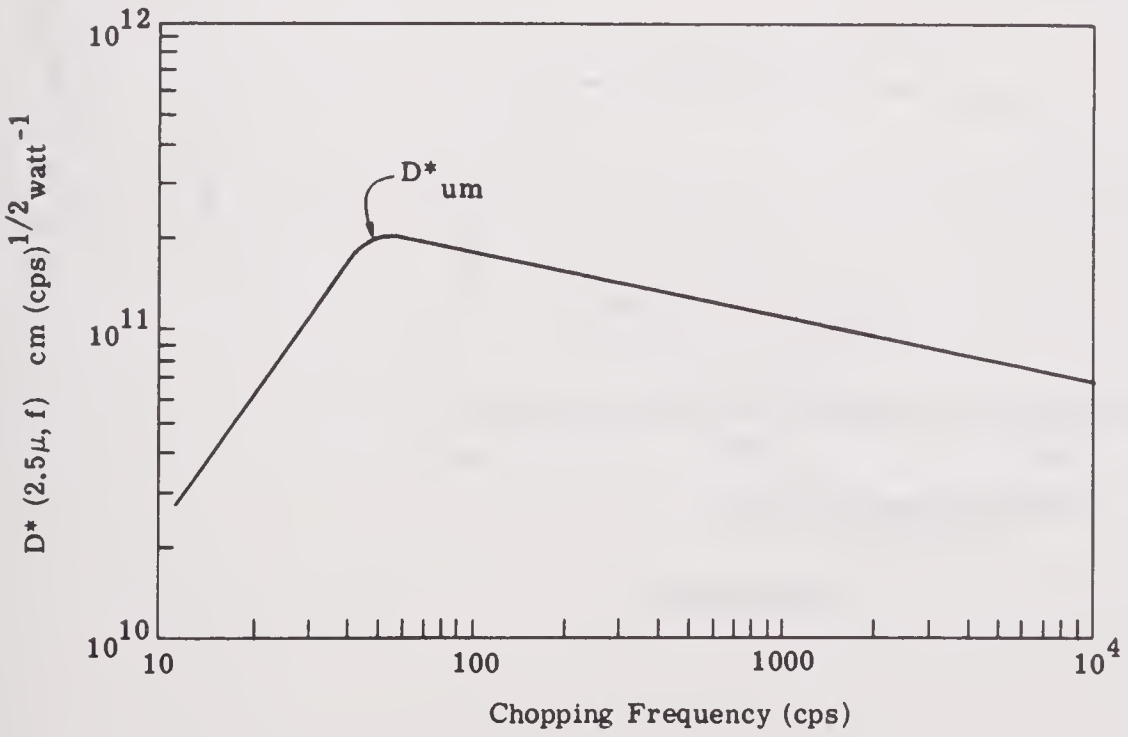
T_B 300°K (background noise insignificant)

* Tradeoffs between parameters are available upon request from the manufacturer (e.g., τ may be shortened at the expense of D*).



T_B	300°K
τ	500-3000 μsec
$\overline{R}$	.5 to 10 MΩ/square
Z	
A	10^{-6} to 10^{0} cm^2
fov	180°
$R_{\lambda_{max}}$	~10^6 $\frac{\text{volt rms}}{\text{watt rms}}$

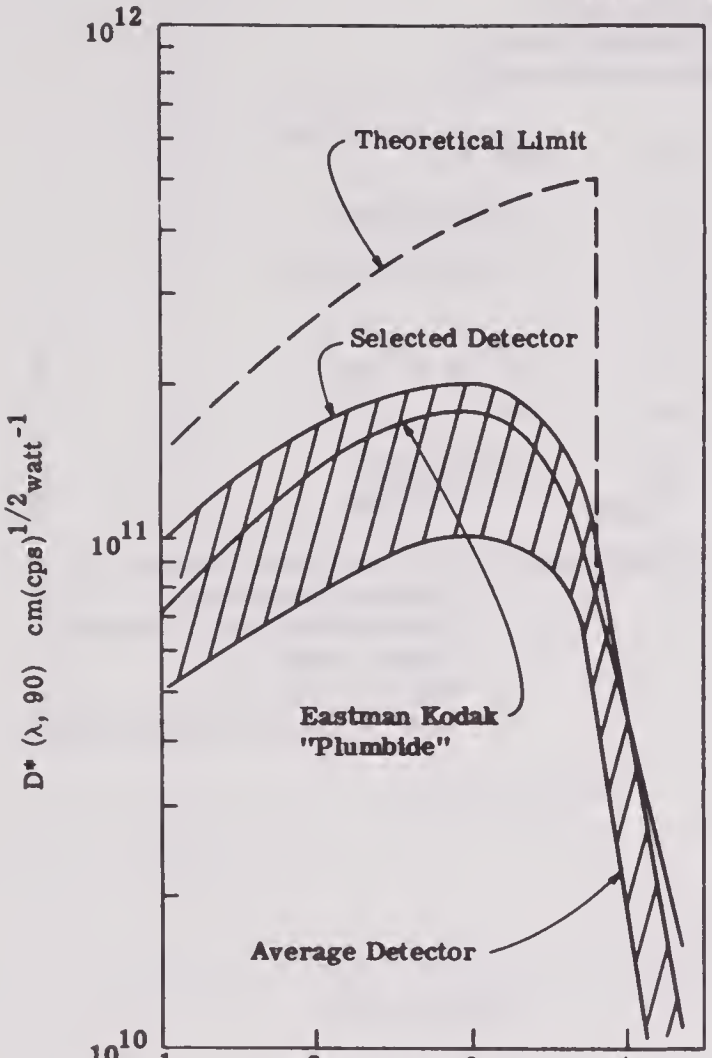
Manufacturers: Infrared Industries Inc.
Eastman Kodak Co.
Electronics Corp. of America
Tupper Trent Co.
Mullard Ltd.
Santa Barbara Research Center



DATA SHEET 5

PbS*

Detector Temperature 78°K



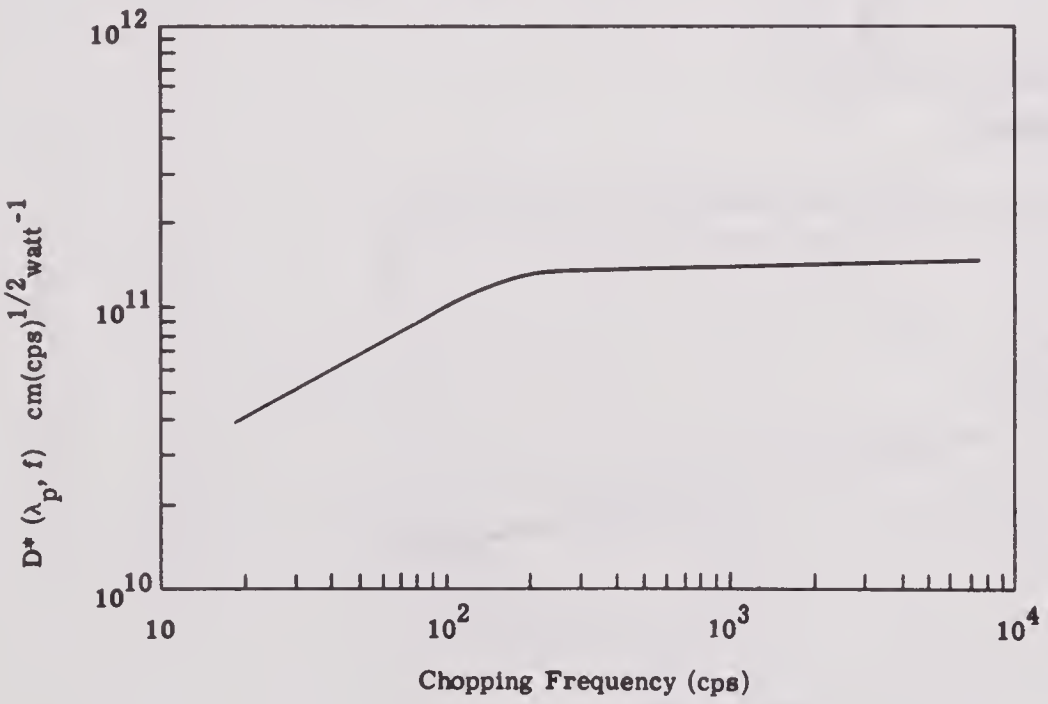
T_B	300°K
τ	1000-5000 μsec
$\overline{R}$	.5-20 MΩ/square
Z	
A	10^{-6} to 10^{0} cm^2
fov	180°
$R_{\lambda_{max}}$	~10^6 $\frac{\text{volt rms}}{\text{watt rms}}$

Operating mode:

Limiting noise:

Manufacturers: Infrared Industries Inc.
Eastman Kodak Co.
Electronics Corp. of America
Tupper Trent Co.
Mullard Ltd.
Santa Barbara Research Center

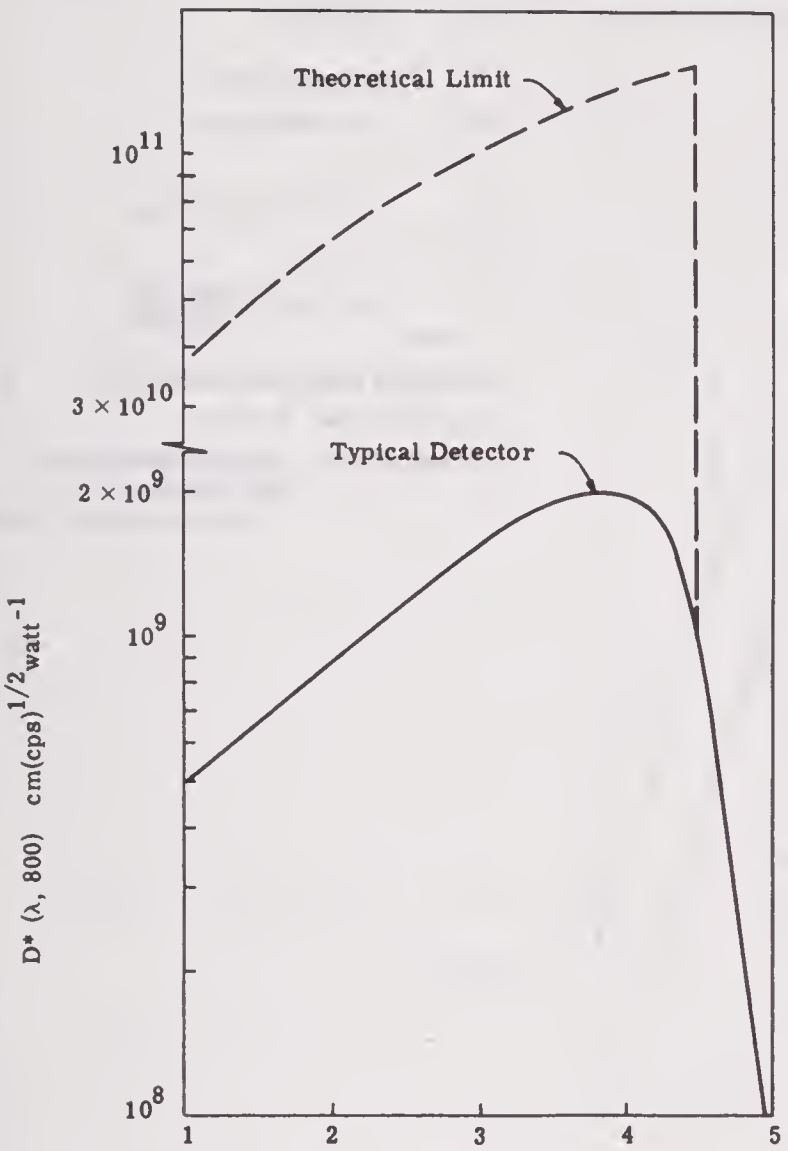
Tradeoffs between parameters are available upon request from the manufacturer (e.g., τ may be decreased at the expense of D).



DATA SHEET 6

PbSe

Detector Temperature 300°K

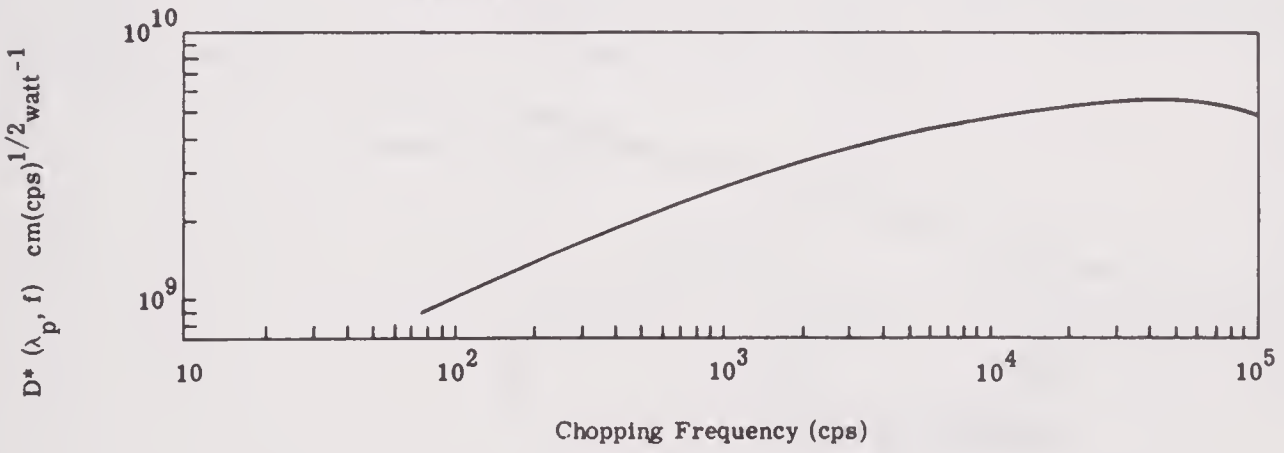


T_B	300°K
τ	1 - 10 μsec
$\bar{R}$	~5 MΩ/square
Z	
A	10^{-6} cm^2 to 10^0 cm
fov	180°
$R_{\lambda_{max}}$	~10^3 $\frac{\text{volt rms}}{\text{watt rms}}$

Operating mode: photoconductive

Limiting noise: 1/f noise

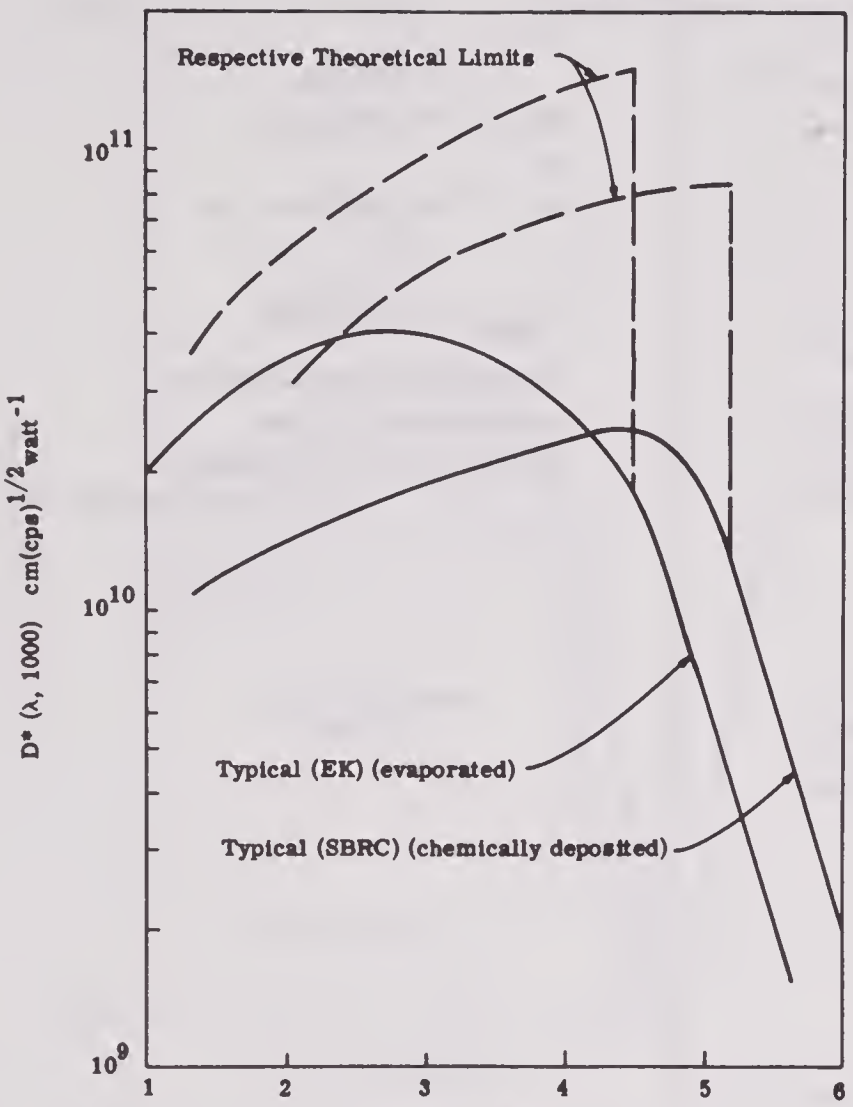
Manufacturers: Eastman Kodak Co.
Santa Barbara Research Center

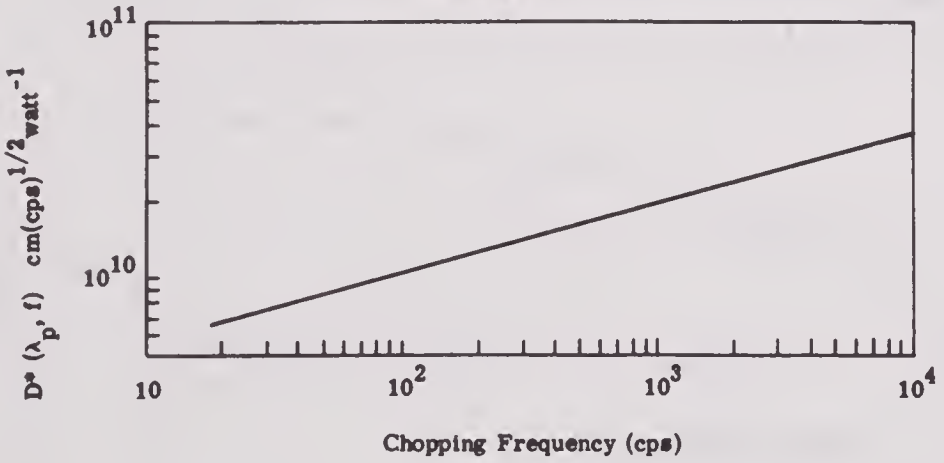


DATA SHEET 7

PbSe

Detector Temperature 195°K





T_B	300°K
τ	10-100 μsec
$\overline{R}$	~50 MΩ/square
Z	
A	10^{-6} cm^2 to 10^{0} cm^2
fov	
$R_{\lambda_{max}}$	10^{5} to 10^{6} $\frac{\text{volt rms}}{\text{watt rms}}$

Operating mode: photoconductive

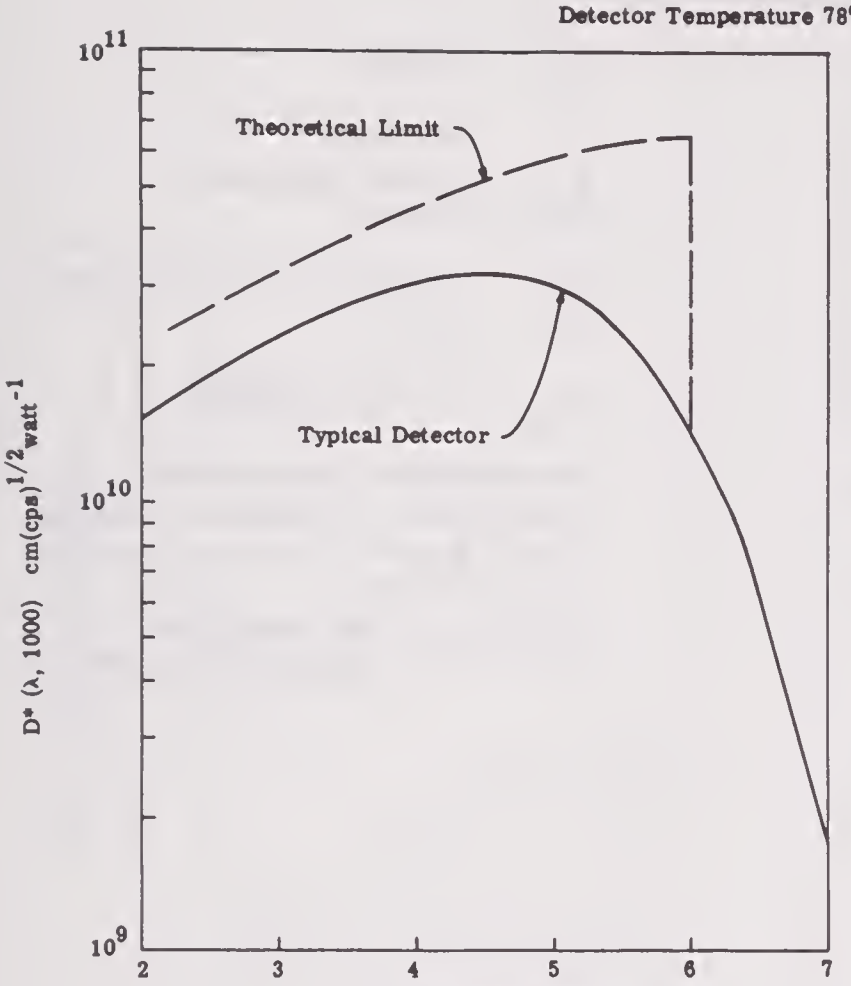
Limiting noise: 1/f noise

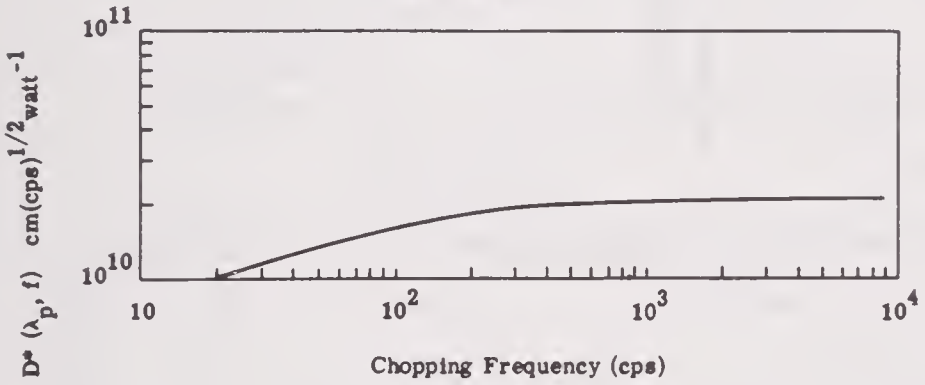
Manufacturers: Eastman Kodak (EK)
Santa Barbara
Research Center (SBRC)

DATA SHEET 8

PbSe

Detector Temperature 78°K





T_B	300°K
τ	10-150 μsec
$\overline{R}$	.5-50 MΩ/square
Z	
A	10^{-6} cm^2 to 10^{0} cm^2
fov	90°
$R_{\lambda_{max}}$	~10^{6} $\frac{\text{volt rms}}{\text{watt rms}}$

Operating mode: photoconductive

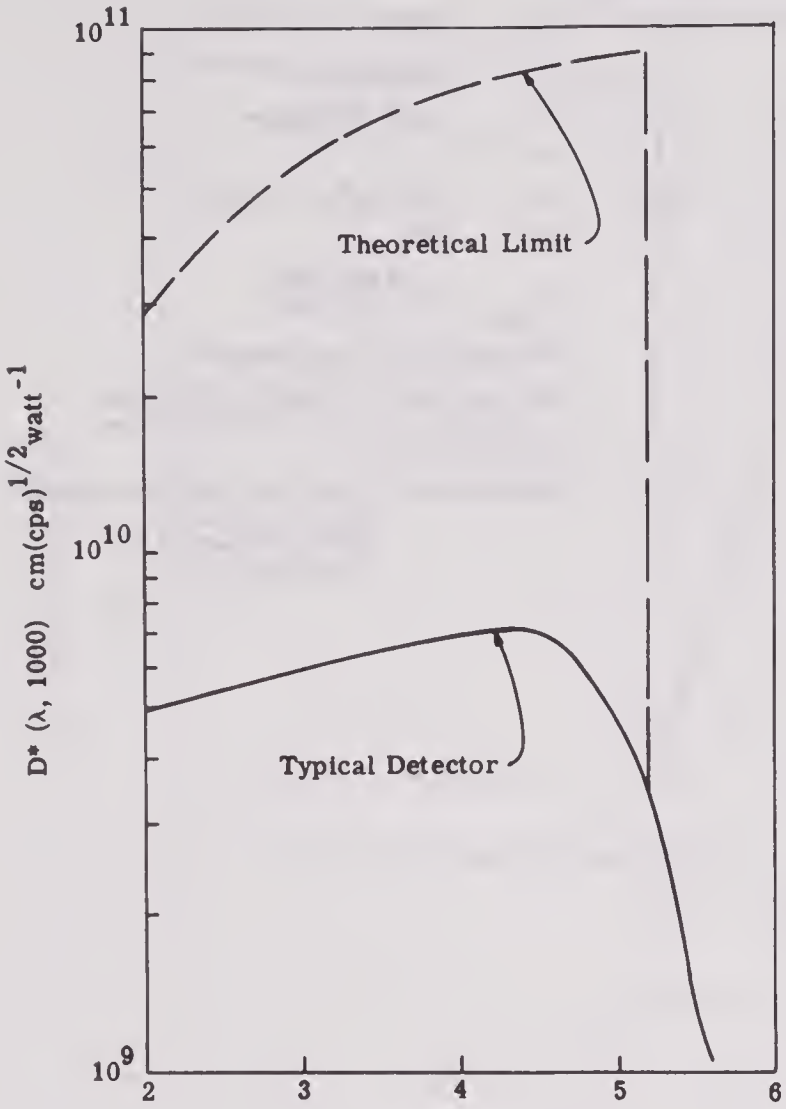
Limiting noise: 1/f noise for f ≤ 400 cps and G-R noise due to lattice vibrations for f ≥ 400 cps.

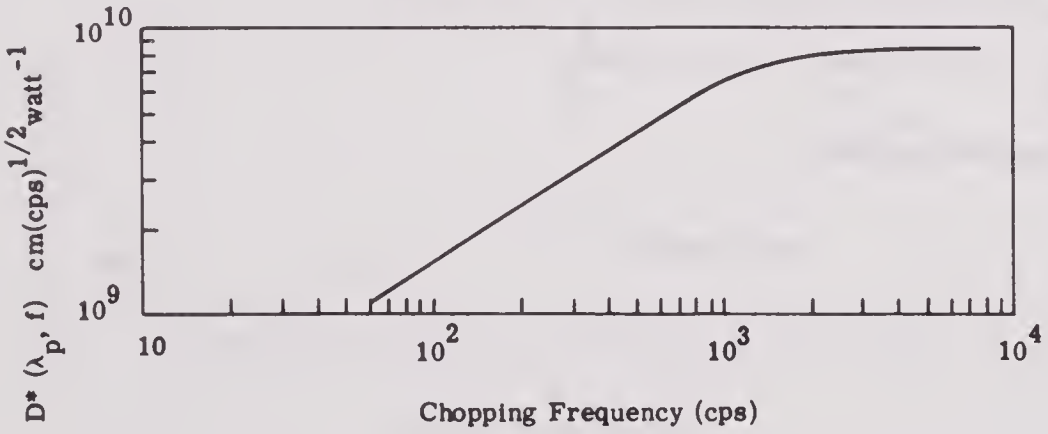
Manufacturers: Santa Barbara Research Center
Eastman Kodak Co.
Infrared Industries Inc.
Librascope Co.

DATA SHEET 9

PbTe

Detector Temperature 78°K





T_B	300°K
τ	10-30 μsec
$\overline{R}$	60-120 MΩ/square
Z	
A	5×10^{-4} cm^2 to 2×10^{-1} cm^2
fov	180°
$R_{\lambda_{max}}$	10^5 to 10^6 $\frac{\text{volt rms}}{\text{watt rms}}$

Operating mode: photoconductive

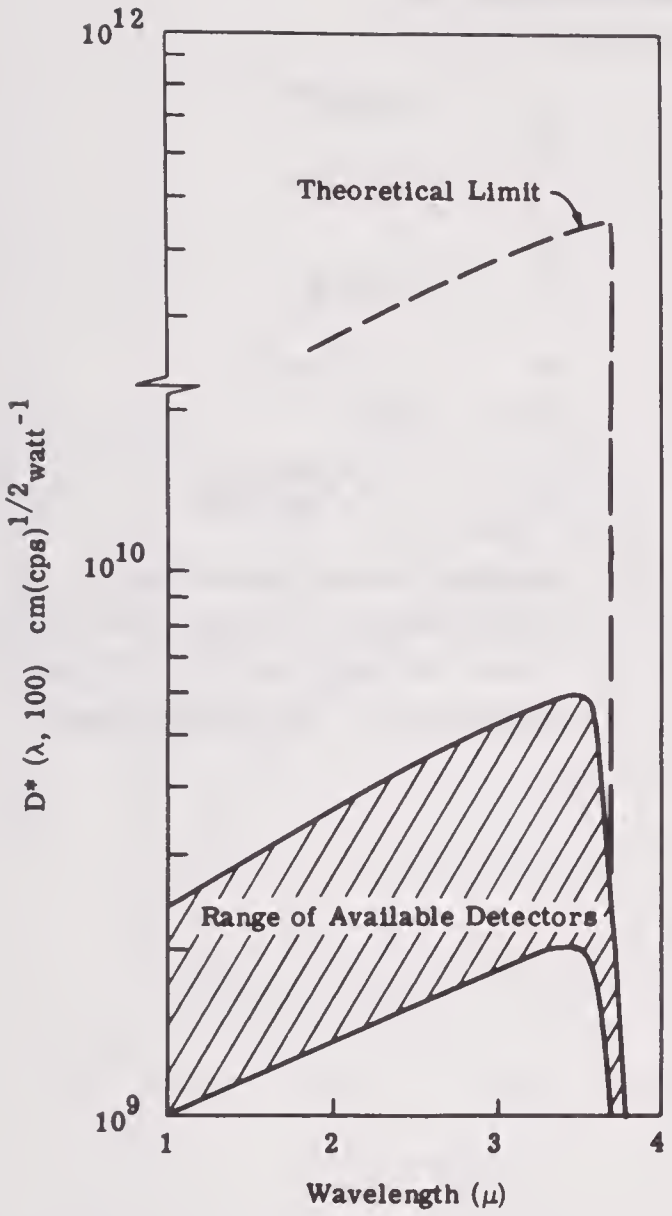
Limiting noise: 1/f noise for f < 1000 cps and G-R noise due to lattice vibrations for f > 1000 cps.

Manufacturer: ITT Laboratories
Minneapolis-Honeywell

DATA SHEET 10

InAs

Detector Temperature 300°K

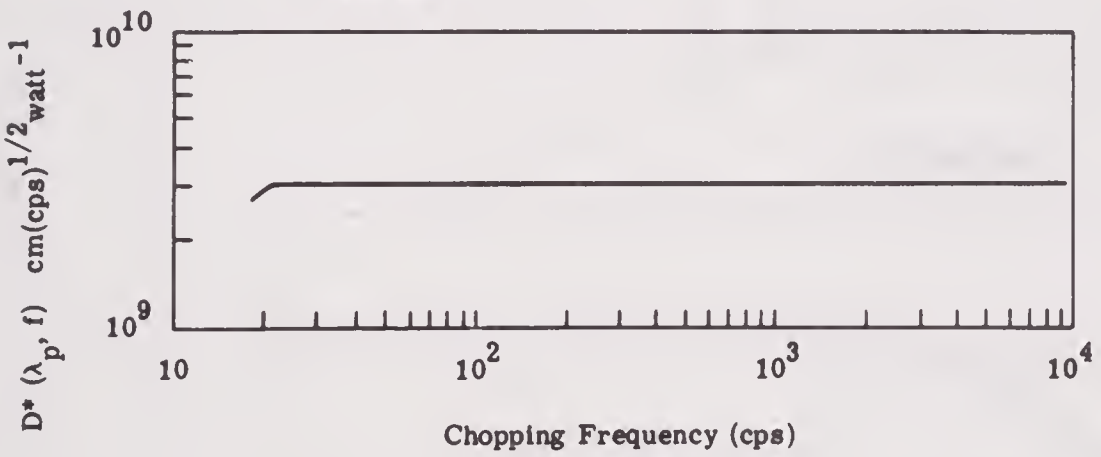


T_B	300°K
τ	<2 μsec
$\overline{R}$	
Z	~20 Ω (requires transformer coupling)
A	10^{-3} to 10^{-1} cm^2
fov	180°
$R_{\lambda_{max}}$	~10 $\frac{\text{volt rms}}{\text{watt rms}}$

Operating mode: photovoltaic

Limiting noise: 1/f noise for f ≤200 cps and Johnson noise for f ≥200 cps.

Manufacturers: Philco Corp. (provides sapphire immersion lens)
Texas Instruments
Electro-Optical Systems

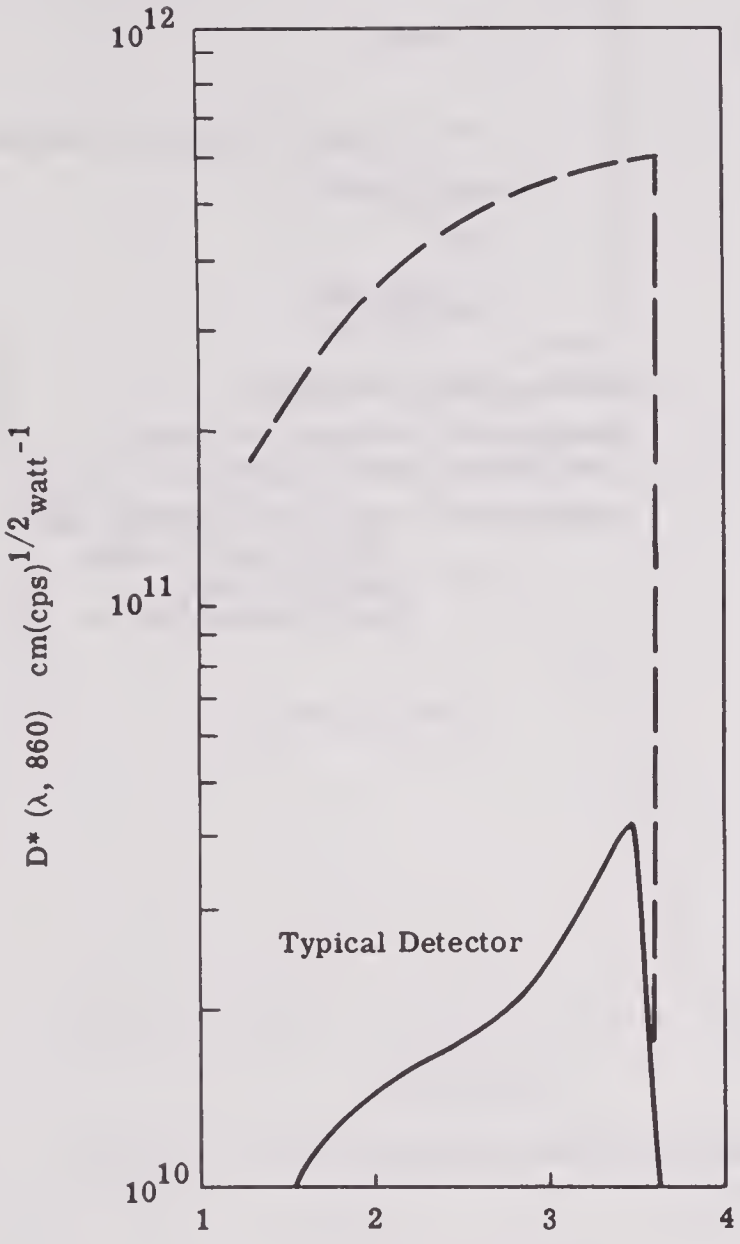


DATA SHEET 11

InAs

Detector Temperature 273°K

(supplied with thermoelectric cooler and an Irtran II immersion lens)

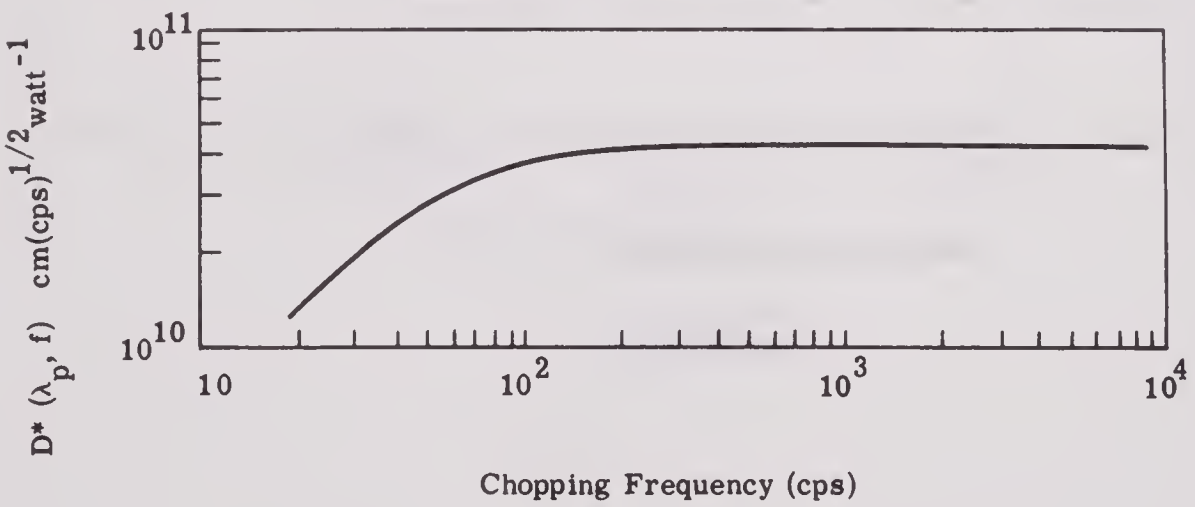


T_B	300°K
τ	<5 μsec
$\overline{R}$	
Z	~150 Ω
A	$\sim 10^{-2}$ cm^2
fov	180°
$R_{\lambda_{max}}$	$\sim 10^3$ $\frac{\text{volt rms}}{\text{watt rms}}$

Operating mode: photovoltaic

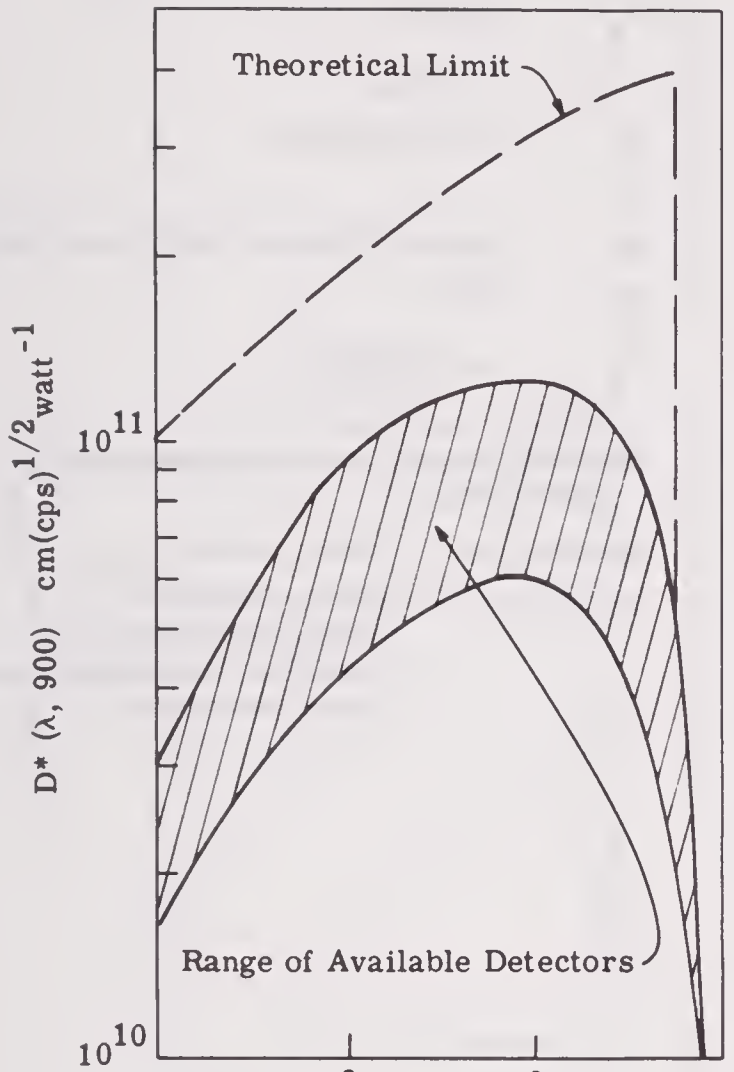
Limiting noise: 1/f noise for f ≤300 cps and Johnson noise for f ≥300 cps.

Manufacturer: Texas Instruments



Te

Detector Temperature 78°K

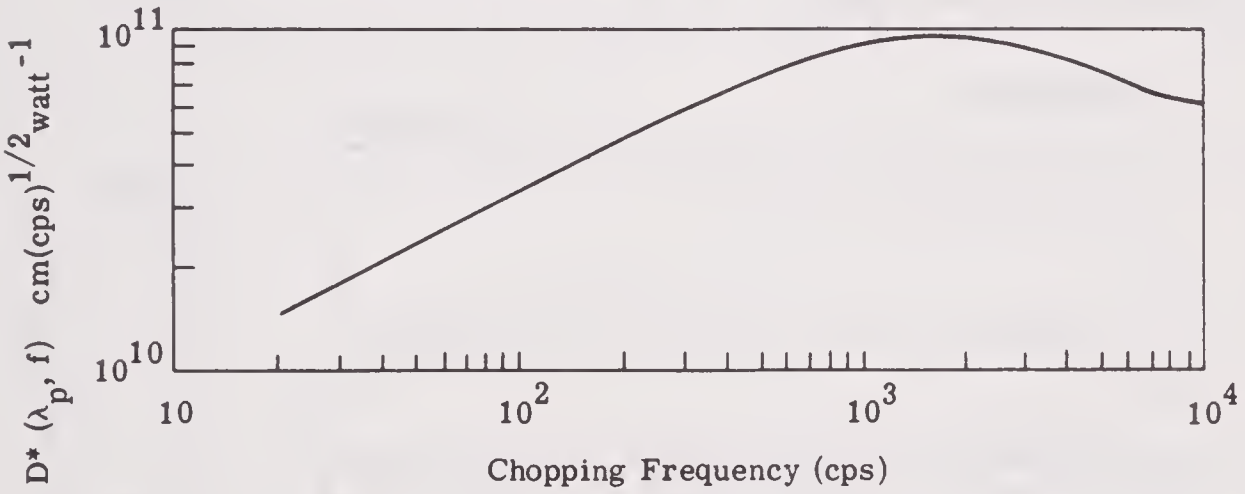


T_B	300°K
τ	60 μsec
$\overline{R}$	1000 ohms/square
Z	
A	$>2 \times 10^{-3}$ cm^2
fov	180°
$R_{\lambda_{max}}$	$\sim 10^4$ $\frac{\text{volt rms}}{\text{watt rms}}$

Operating mode: photoconductive

Limiting noise: 1/f noise for f <1000 cps and G-R noise due to lattice vibrations (phonons) for f >1000 cps.

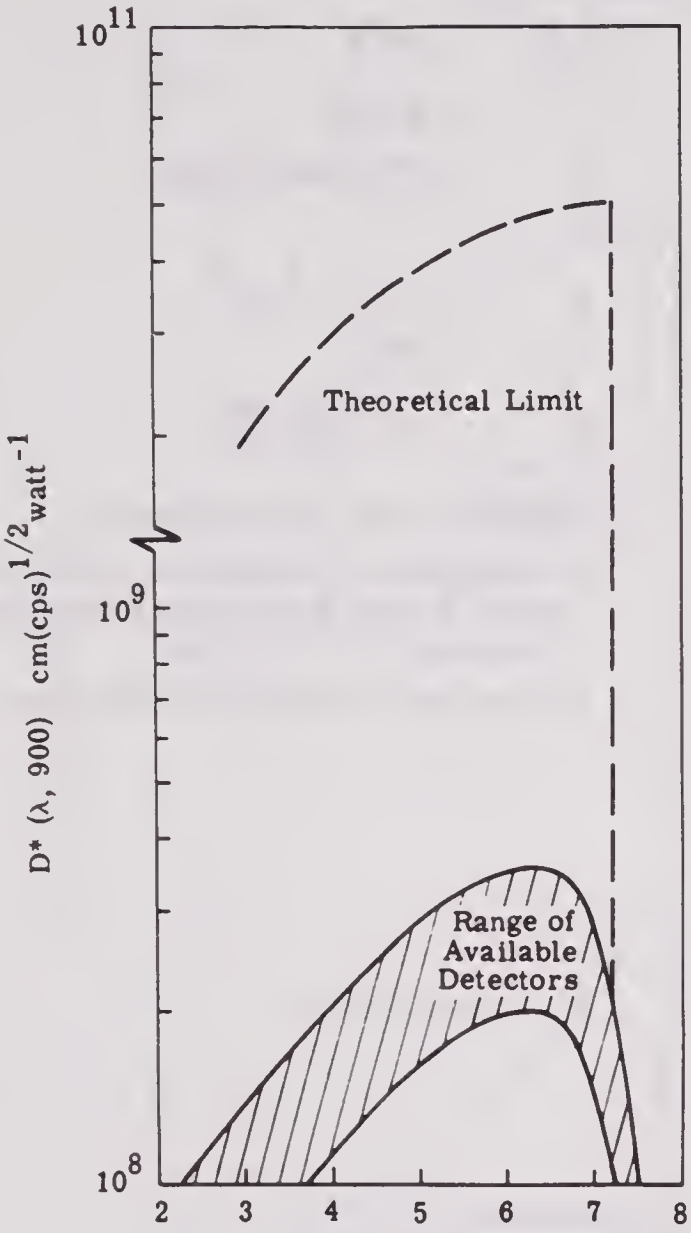
Manufacturer: Minneapolis-Honeywell



DATA SHEET 13

InSb

Detector Temperature 300°K

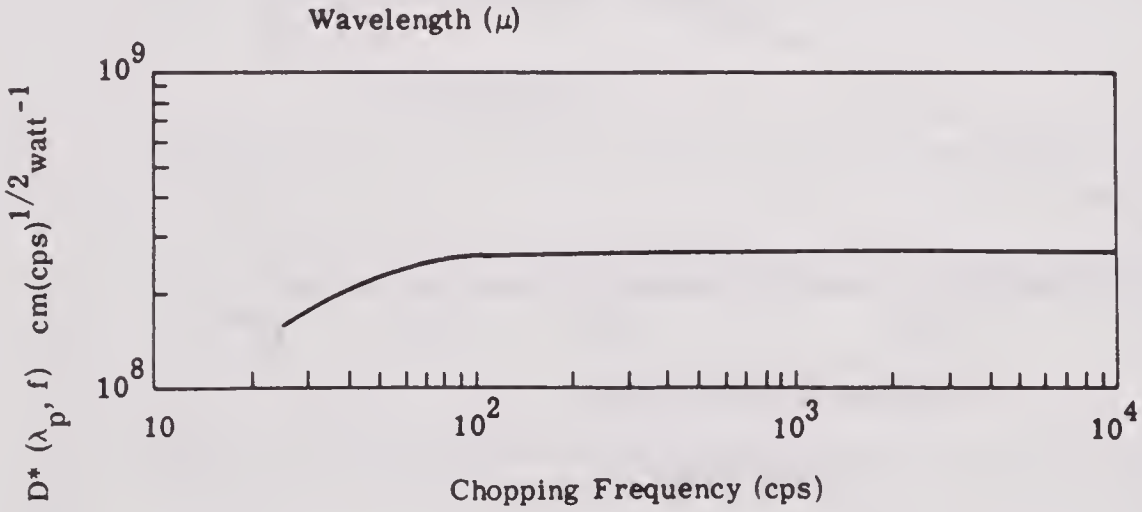


T_B	300°K
τ	< .2 μsec
$\overline{R}$	~15 ohms/square
Z	
A	(0.05 × 0.05) cm^2 to (0.2 × 0.5) cm^2
fov	180°
$R_{\lambda_{max}}$	~5 $\frac{\text{volt rms}}{\text{watt rms}}$

Operating mode: photoelectromagnetic or photoconductive

Limiting noise: Johnson noise

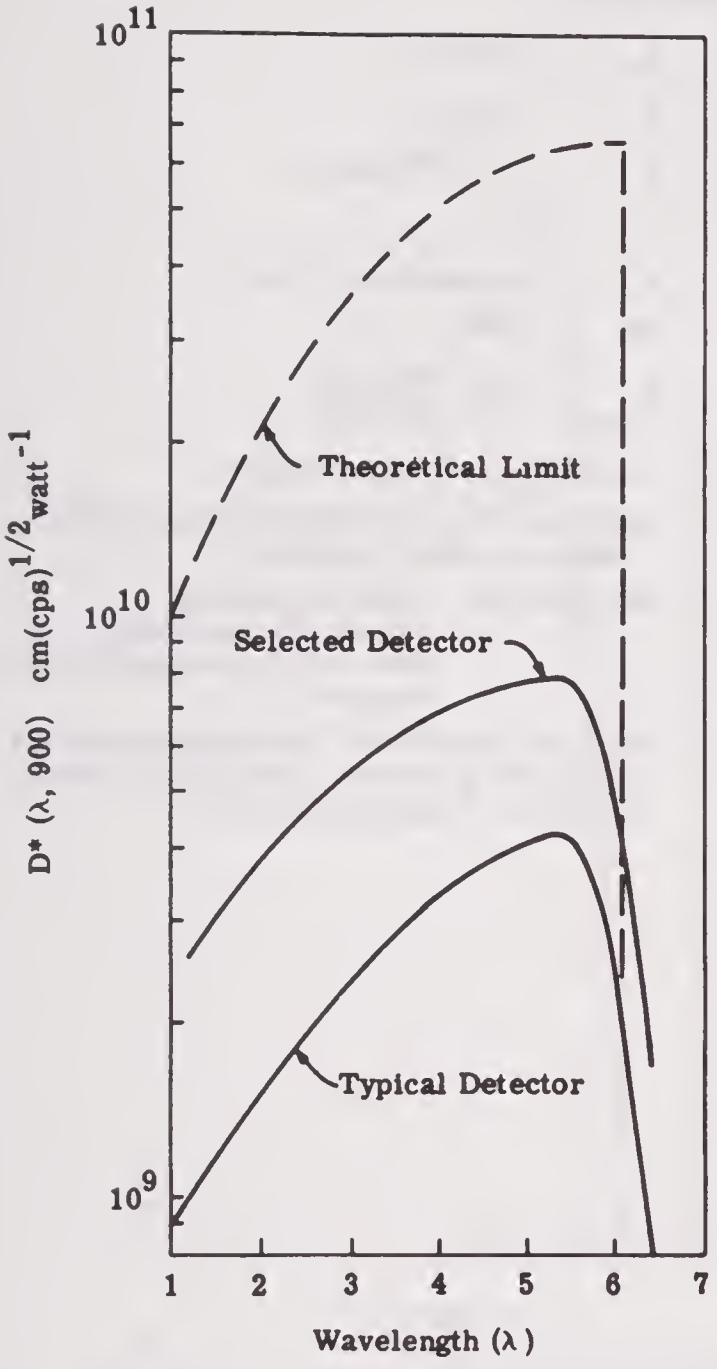
Manufacturers: Block Associates
Minneapolis-Honeywell
Radiation Electronics Corp.
Texas Instruments



DATA SHEET 14

InSb

Detector Temperature 195°K

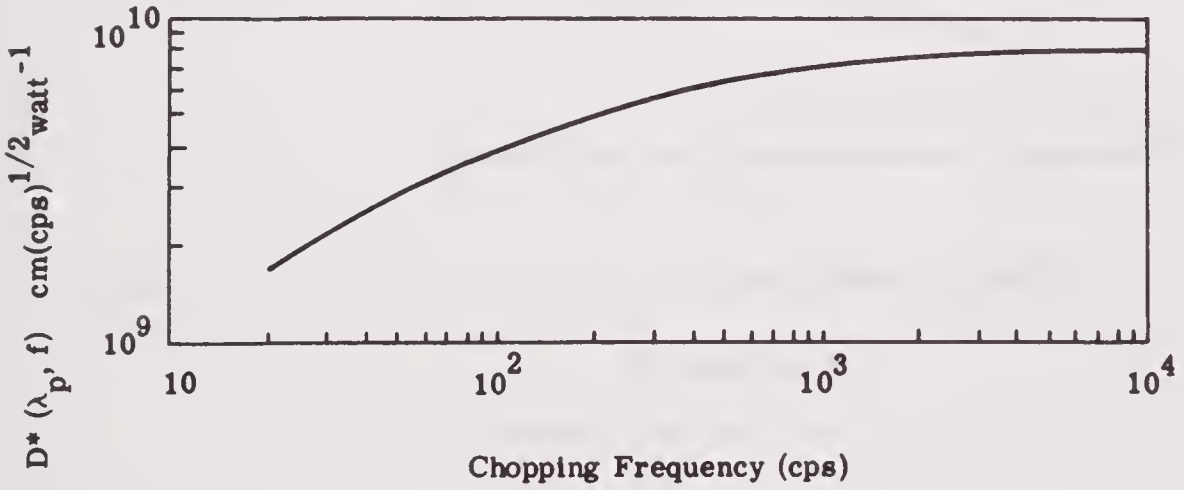


T_B	300°K
τ	<1 μsec
$\overline{R}$	~100 ohms/square
Z	
A	(.05 × .01) cm^2 to (0.5 × 0.5) cm^2
fov	180°
$R_{\lambda_{max}}$	~10^2 $\frac{\text{volt rms}}{\text{watt rms}}$

Operating mode: photoconductive

Limiting noise: 1/f noise plus G-R noise due to lattice vibrations

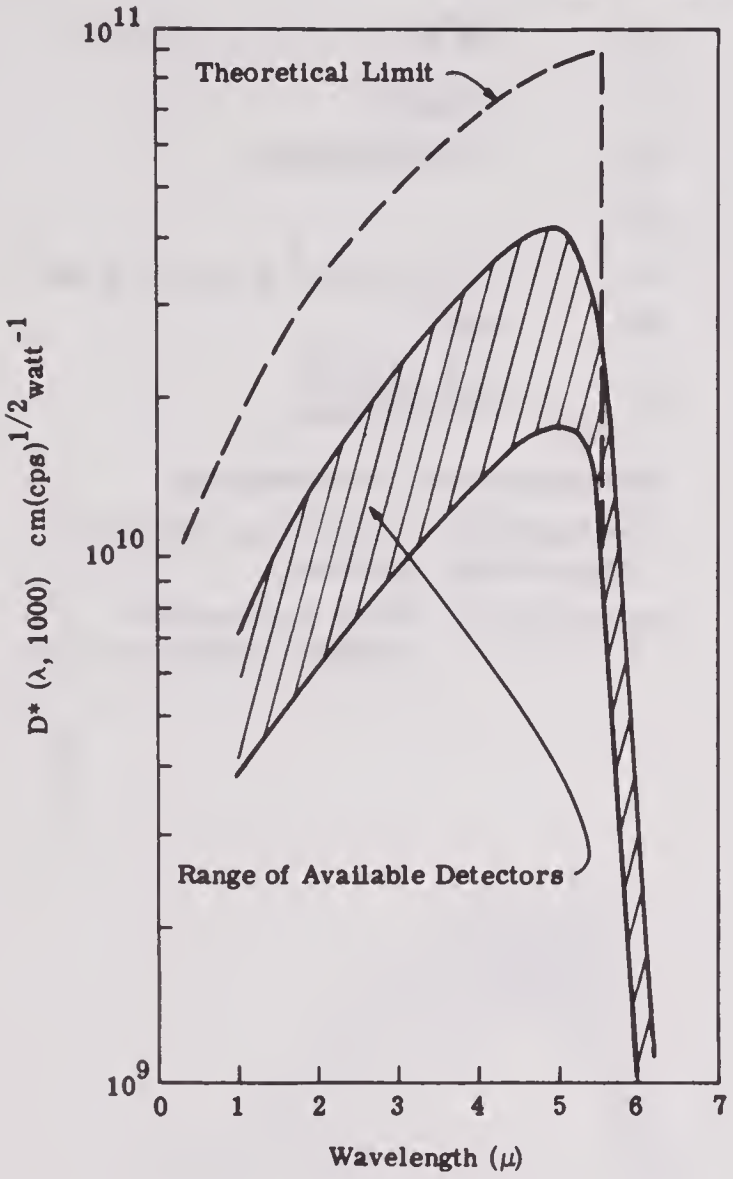
Manufacturers: Texas Instruments
Radiation Electronics Corp.



DATA SHEET 15

InSb

Detector Temperature 78°K



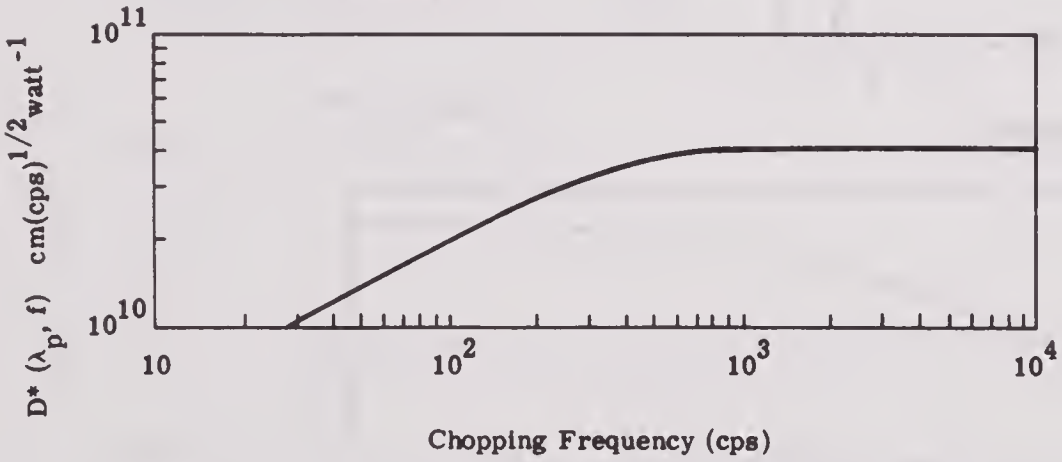
T_B	300°K
τ	~1 μsec
$\overline{R}$	1-20 KΩ/square
Z	
A*	as small as 10^{-3} cm^2
fov	180°
$R_{\lambda_{max}}$	~10^4 $\frac{\text{volt rms}}{\text{watt rms}}$

Operating mode: photoconductive

Limiting noise: G-R noise due to background photons to lattice vibrations

Manufacturers: Texas Instruments
Minneapolis-Honeywell
Santa Barbara Research Center
Raytheon

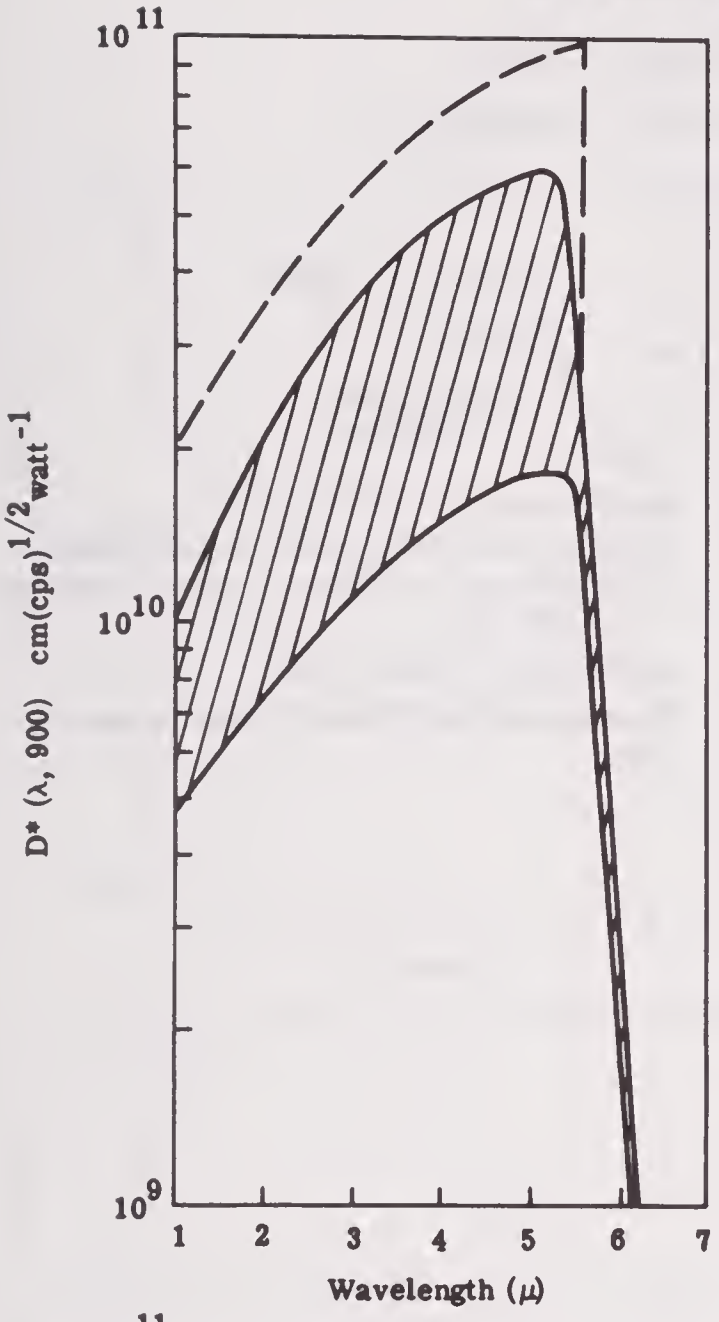
*Areas as small as 10^{-5} have been constructed by the Chicago Midway Laboratories (now the Laboratory of Applied Science)



DATA SHEET 16

InSb

Detector Temperature 78°K

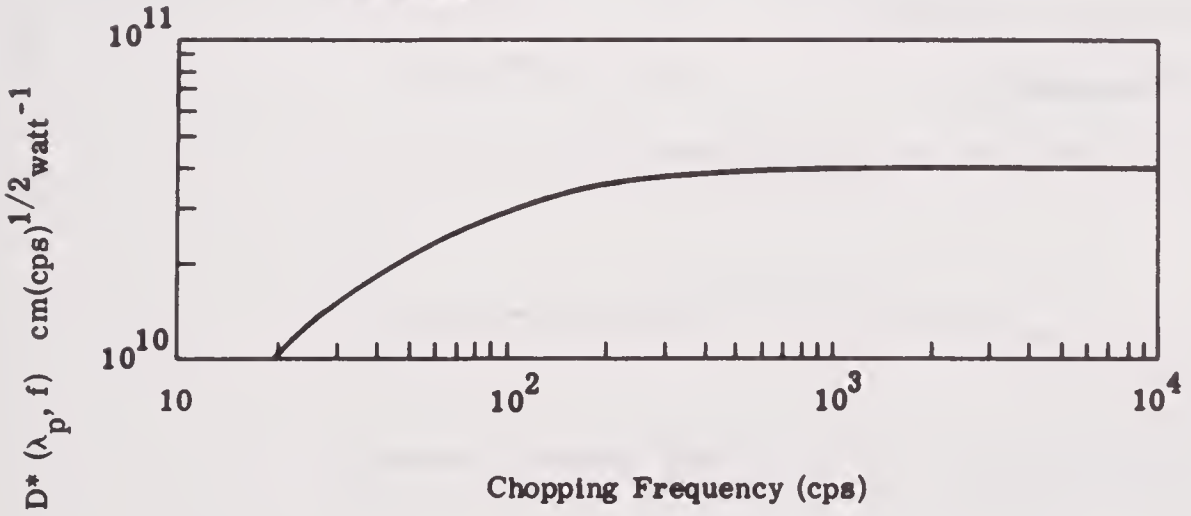


T_B	300°K
τ	<1 μsec
$\overline{R}$	
Z	200 to 10 MΩ
A	10^{-3} cm^2 to .3 cm^2
fov	180°
$R_{\lambda_{max}}$	10^4 to 10^5 $\frac{\text{volt rms}}{\text{watt rms}}$

Operating mode: photovoltaic

Limiting noise: 1/f noise for f <300 cps and shot noise for f >300 cps arising from the generation of minority carriers due to photon and phonon excitation.

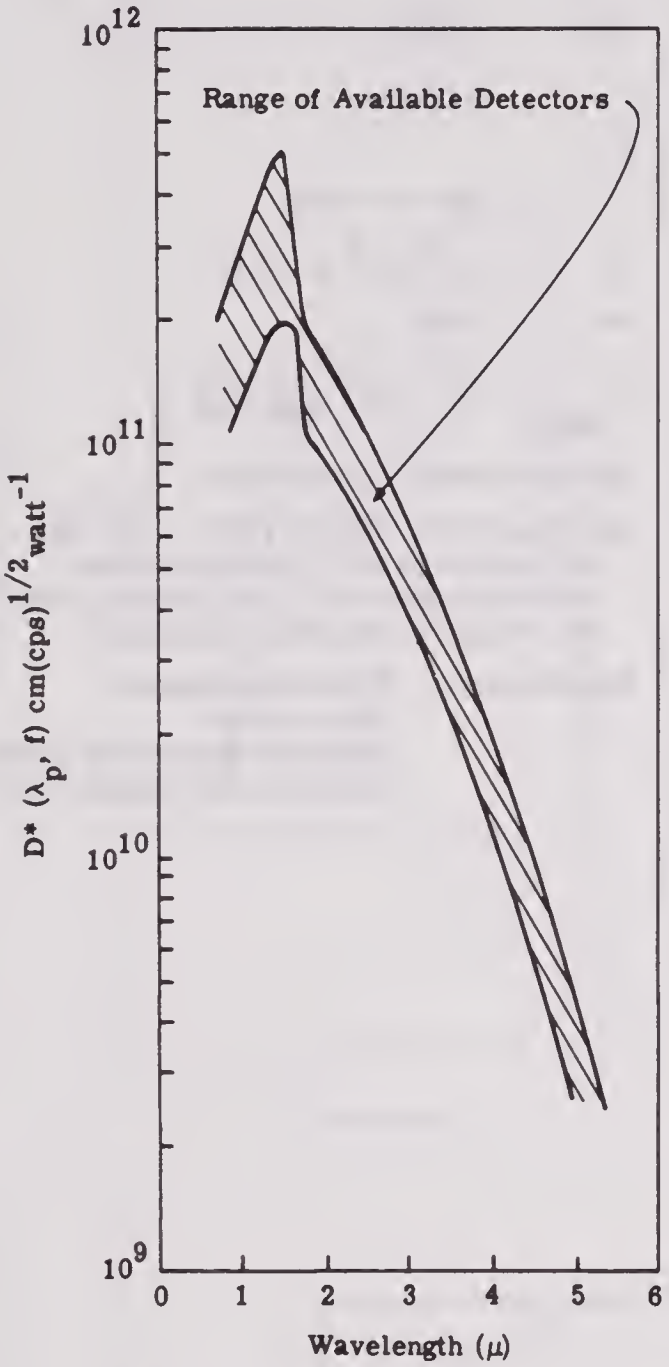
Manufacturers: Texas Instruments
Philco Corp.
Networks Electronics Corp.
Radiation Electronics Corp.



DATA SHEET 17

Ge:AuSb* (n-type)

Detector Temperature 78°K



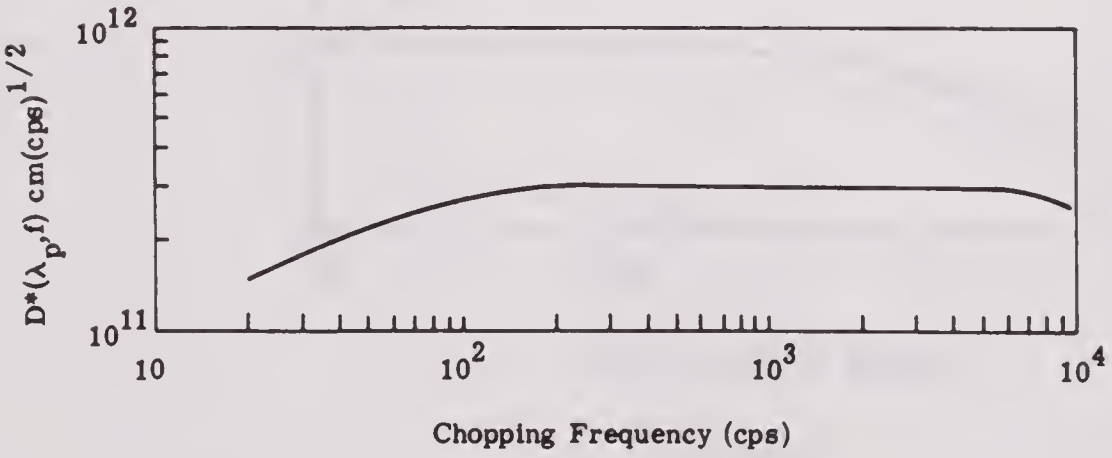
T_B	300°K
τ	20-800 μsec
$\overline{R}$	2-20 MΩ
Z	
A	10^{-3} - 10^{-1} cm^2
fov	~110°
$R_{\lambda_{max}}$	~10^6 $\frac{\text{volt rms}}{\text{watt rms}}$

Operating mode: photoconductive

Limiting noise: 1/f noise for f <200 cps and G-R noise due to lattice excitations (phonons) for f >200 cps.

Manufacturer: Philco Corp.

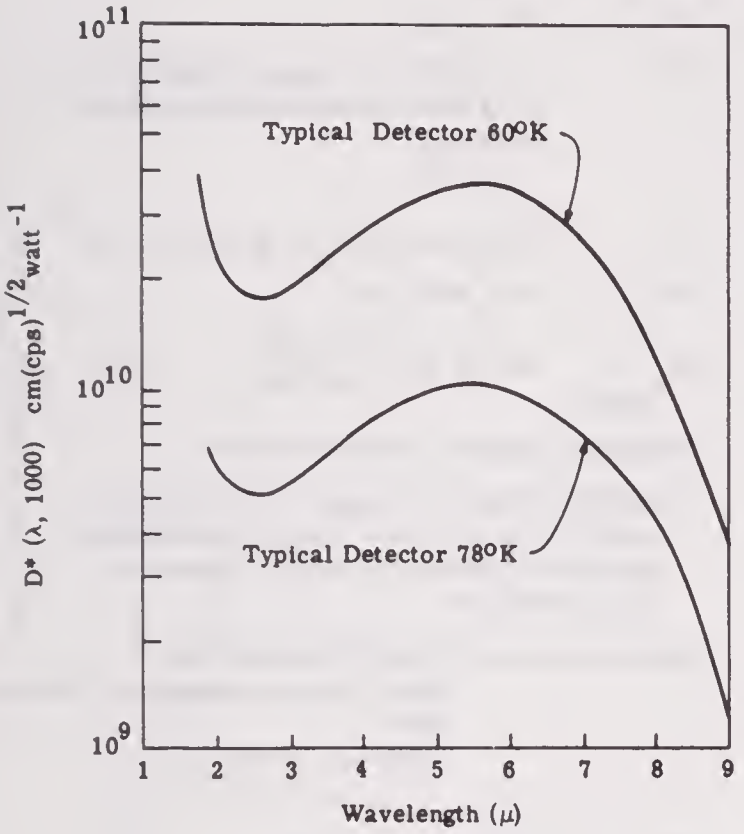
*Manufacturer will provide a lens for optical gain.



DATA SHEET 18

Ge:AuSb (p-type)

Detector Temperature 78°K or 60°K

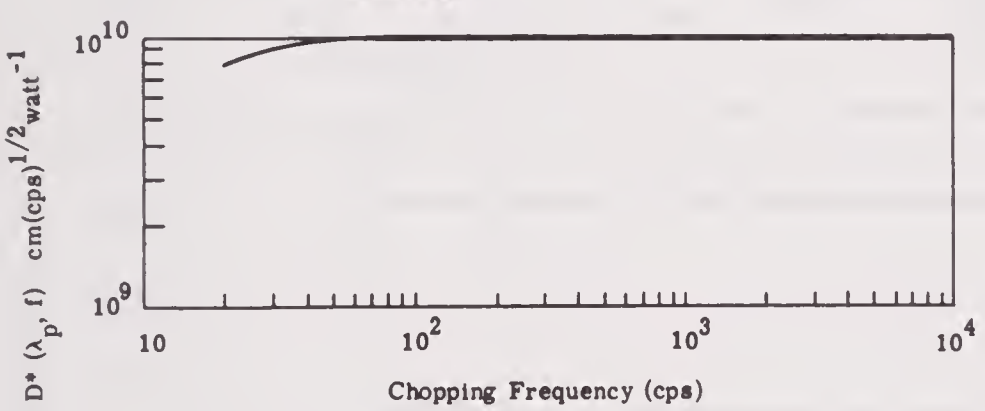


T_B	300°K
τ	<1 μsec
$\overline{R}$	.1-5.0 MΩ/square at 78°K
Z	
A	(.05 × .05) cm^2 to (0.5 × 0.5) cm^2
fov	60°
$R_{\lambda_{max}}$	~10^4 $\frac{\text{volt rms}}{\text{watt rms}}$ at 78°K

Operating mode: photoconductive

Limiting noise: 1/f noise for f <100 cps and G-R noise due to lattice excitations (phonons) for f >100 cps. For detector temperatures ≤60°K the G-R noise has a significant photon noise component.

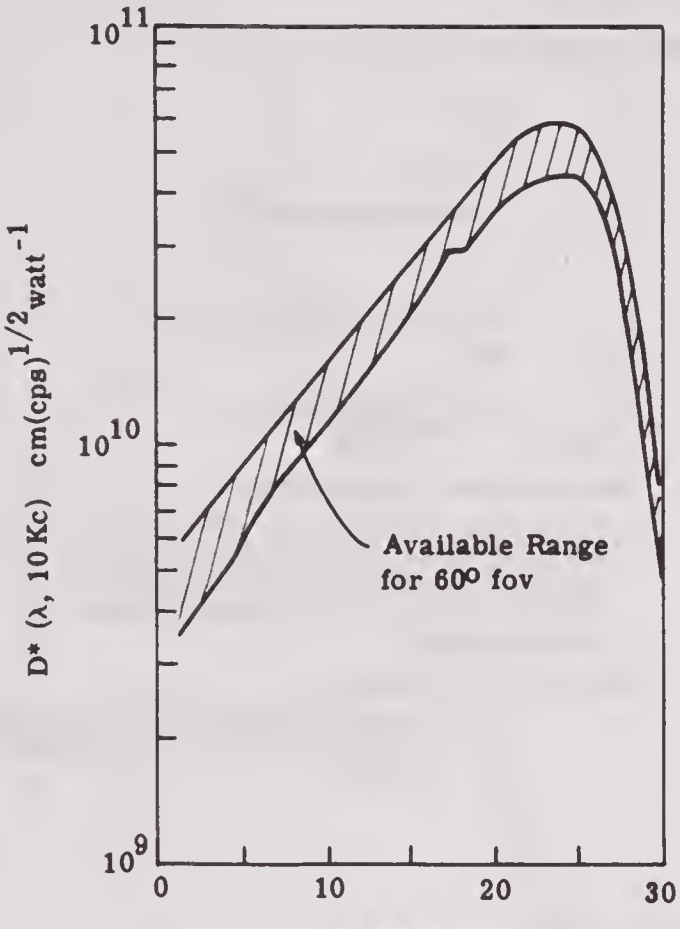
Manufacturers: Philco Corp.
Westinghouse Research Laboratories
Santa Barbara Research Laboratory
Raytheon

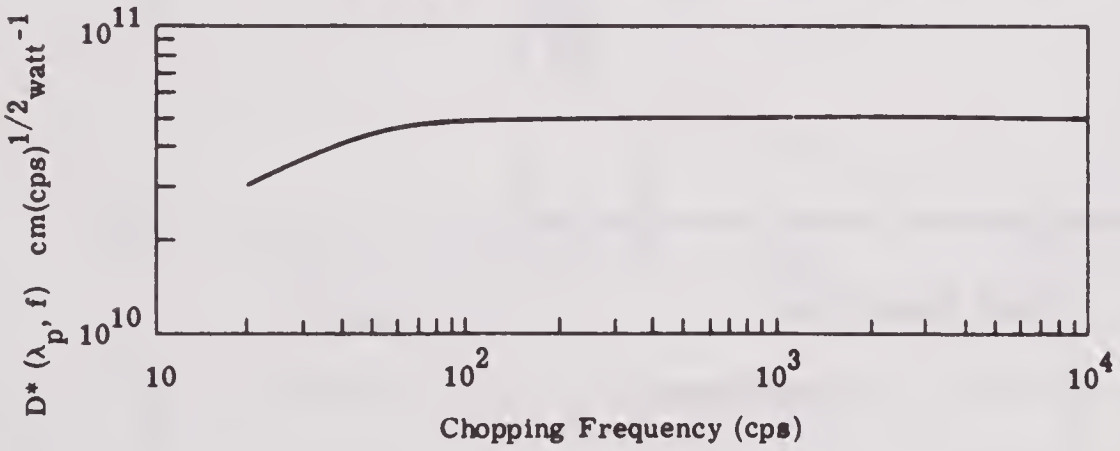


DATA SHEET 19

Ge:Cu

Detector Temperature ~ 4°K





T_B	300°K
τ	<1 μsec
$\overline{R}$	.02 to 20 MΩ square (depends on bias voltage and background temperature)
Z	
A	(.05 × .05) cm^2 to (0.5 × 0.5) cm^2
fov	26°, 60°, 90°
$R_{\lambda_{max}}$	10^4 to 10^6 $\frac{\text{volt rms}}{\text{watt rms}}$

Operating mode: photoconductive

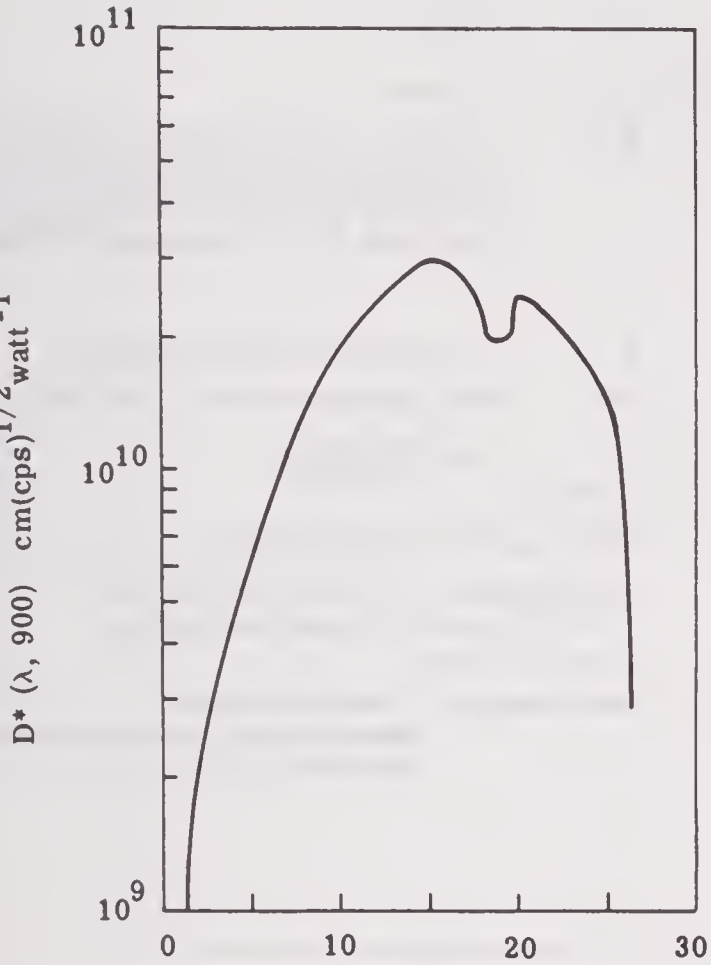
Limiting noise: 1/f noise for f <100 cps and G-R noise due to background photons and due to lattice vibrations (phonons) for f >100 cps.

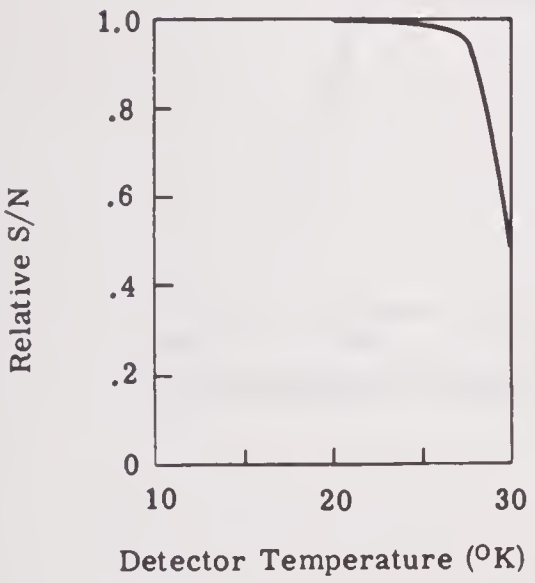
Manufacturers: Texas Instruments
Santa Barbara Research Center
RCA
Eastman Kodak Co.

DATA SHEET 20

Ge:Cd

Detector Temperature 20°K





T_B 300°K

τ <1 μsec

$\overline{R}$

Z

A $(.05 \times .05)$ cm^2 to (0.5×0.5) cm^2

fov 60° (other sizes are available)

R_{λ} max

Operating mode: photoconductive

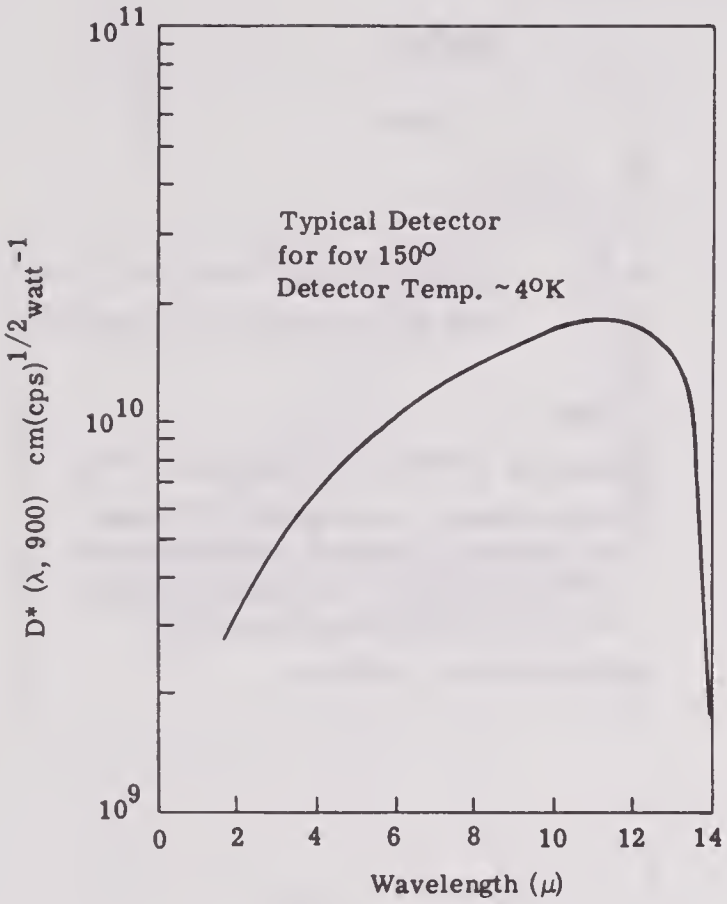
Limiting noise: presumably 1/f noise for smaller chopping frequencies and G-R noise due to background photons for larger chopping frequencies.

Manufacturer: Raytheon

DATA SHEET 21

Ge:Hg

Detector Temperature ~4°K

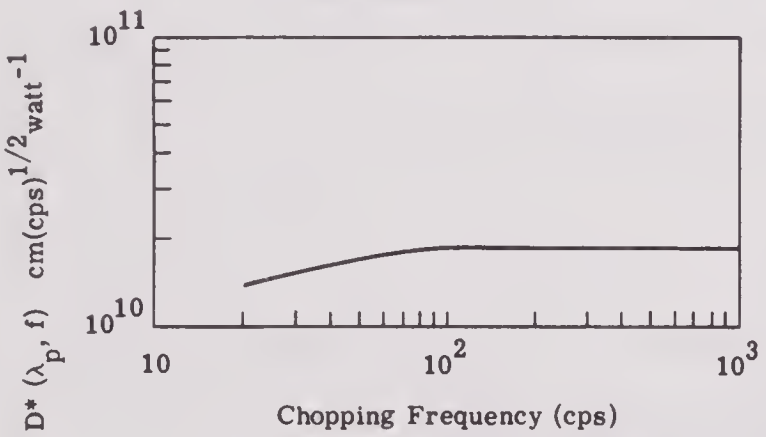


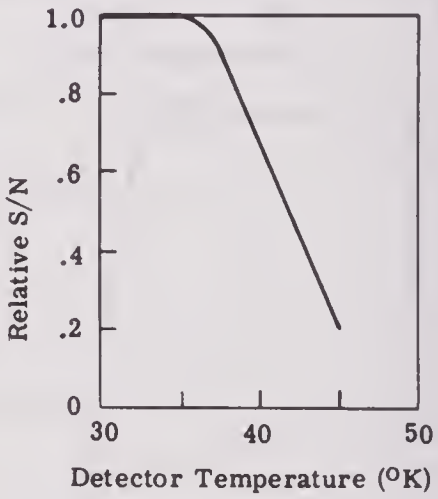
T_B	300°K
τ	<1 μsec
$\overline{R}$	a function of the amount of background radiation when the detector is photon noise limited. For the conditions specified here, $\overline{R}$ is in the range 40-400 kΩ.
Z	
A	$(.05 \times .05)$ cm^2 to (0.5×0.5) cm^2
fov	≈150° (can range from ~8° to 180°)
$R_{\lambda_{max}}$	$\sim 10^6 \frac{\text{volt rms}}{\text{watt rms}}$

Operating mode: photoconductive

Limiting noise: 1/f for f ≤100 cps and G-R noise due to background photons for f >100 cps

Manufacturers: Texas Instruments
Santa Barbara Research Center
Raytheon

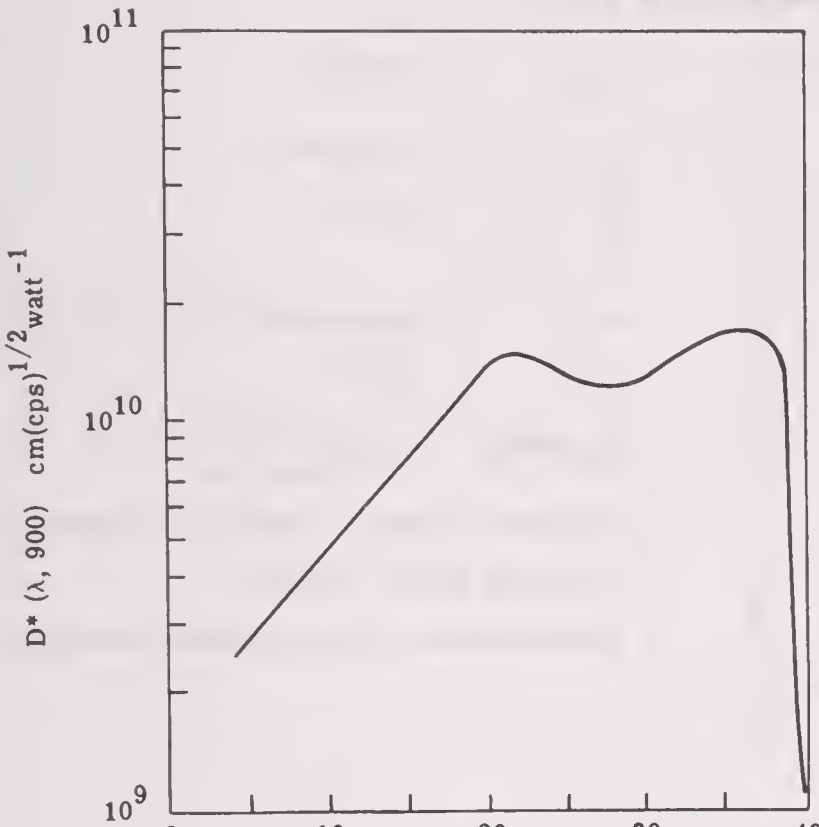




DATA SHEET 22

Ge:Zn

Detector Temperature 4°K



T_B	300°K
τ	< .01 μsec
$\bar{R}$	.2 to 20 MΩ/square
Z	
A	(.05 × .05) cm^2 to (.5 × .5) cm^2
fov	60°
$R_{\lambda_{max}}$	

Operating mode: photoconductive

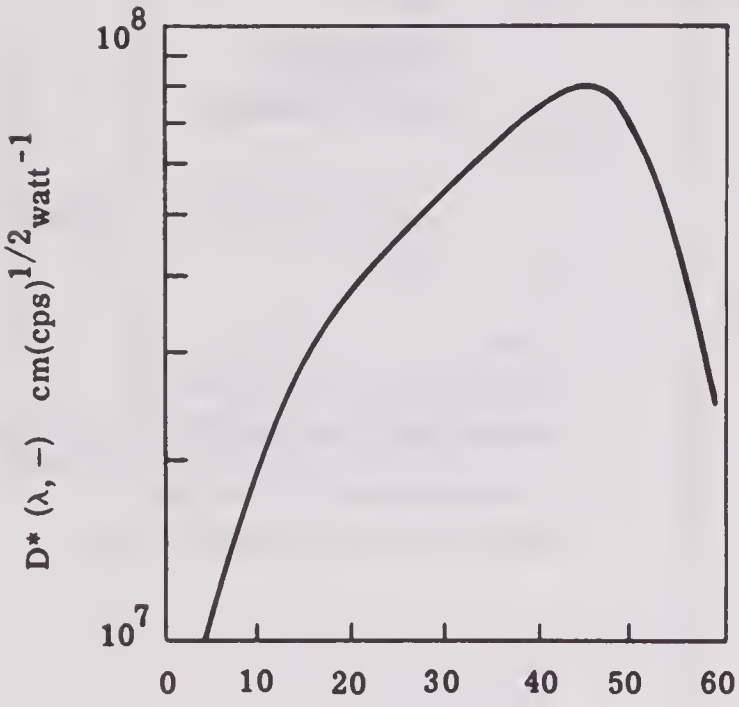
Limiting noise: presumably 1/f noise limited for f < 100 cps and G-R noise due to background photons for f > 100 cps.

Manufacturer: Perkin-Elmer Corp.

DATA SHEET 23

HgTe [5% ZnTe, 5% CdTe]

Detector Temperature 500°K



T_B	300°K
τ	≈2 μsec
$\overline{R}$	≈5 Ω
Z	
A	minimum (.05 × .05) cm^2
fov	180°
R(500°K)	≈.05 $\frac{\text{volt rms}}{\text{watt rms}}$

Operating mode: photoelectromagnetic

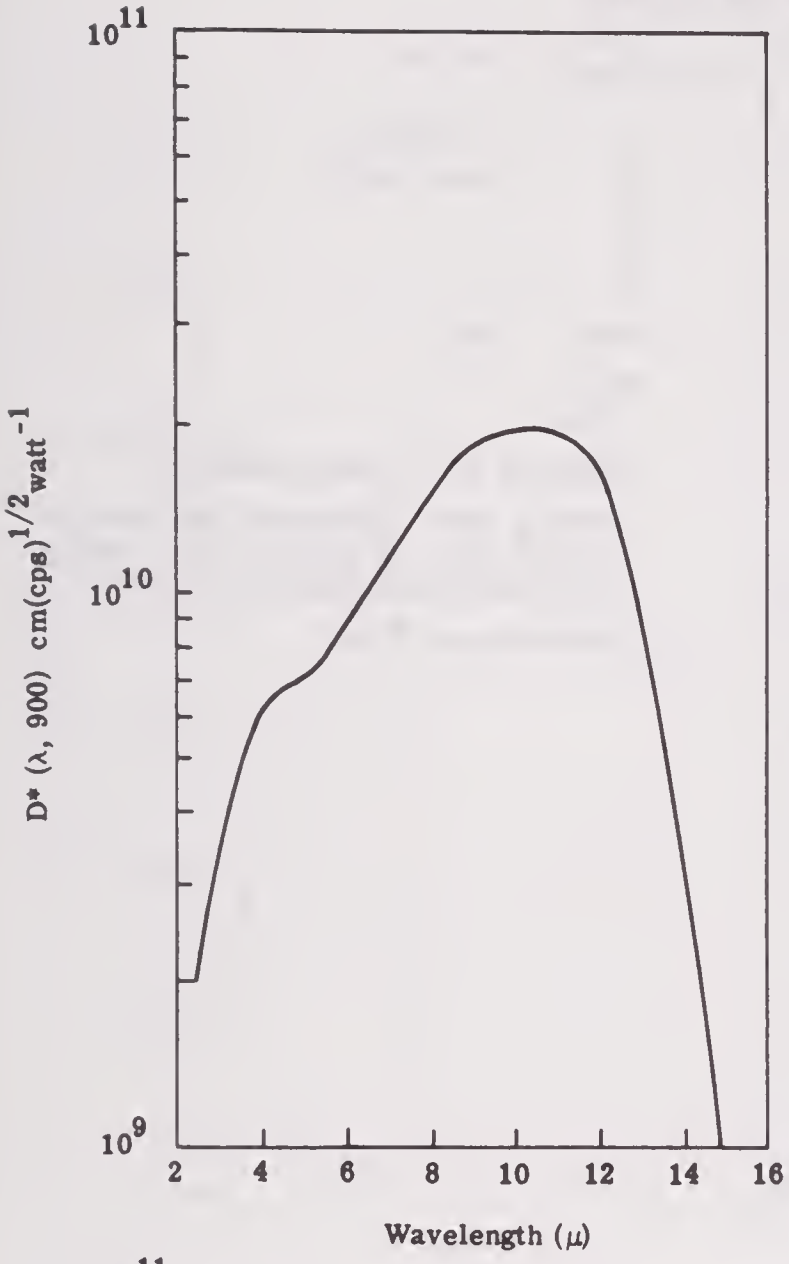
Limiting noise: Johnson

Manufacturer: Minneapolis-Honeywell

DATA SHEET 24

Ge-Si:Zn

Detector Temperature 50°K

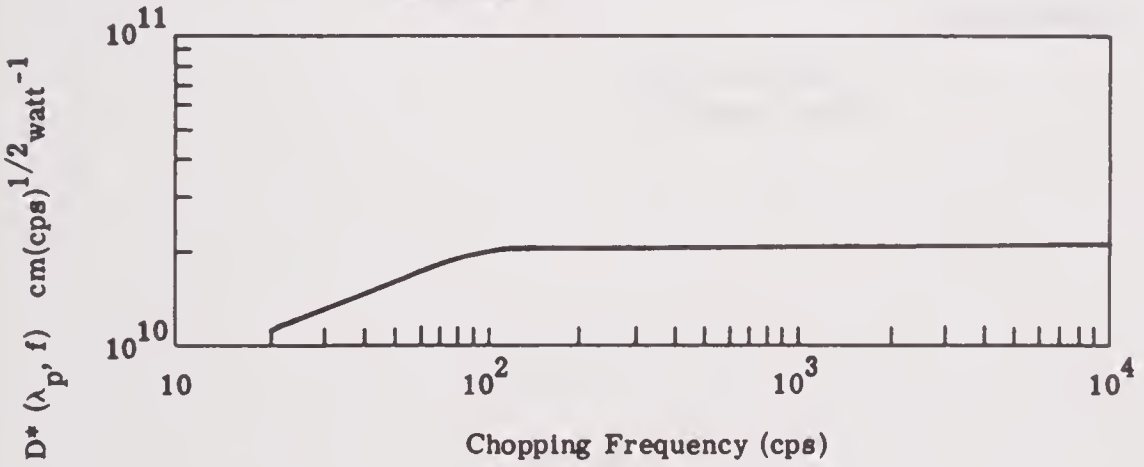


T_B	300°K
τ	<1 μsec
$\overline{R}$	~3×10^6 Ω
Z	
A	
fov	20°
$R_{\lambda max}$	

Operating mode: photoconductive

Limiting noise: for chopping frequencies higher than 100 cps G-R noise due to photons and lattice excitations (phonons) predominate. 1/f noise becomes significant for f ≤ 100 cps.

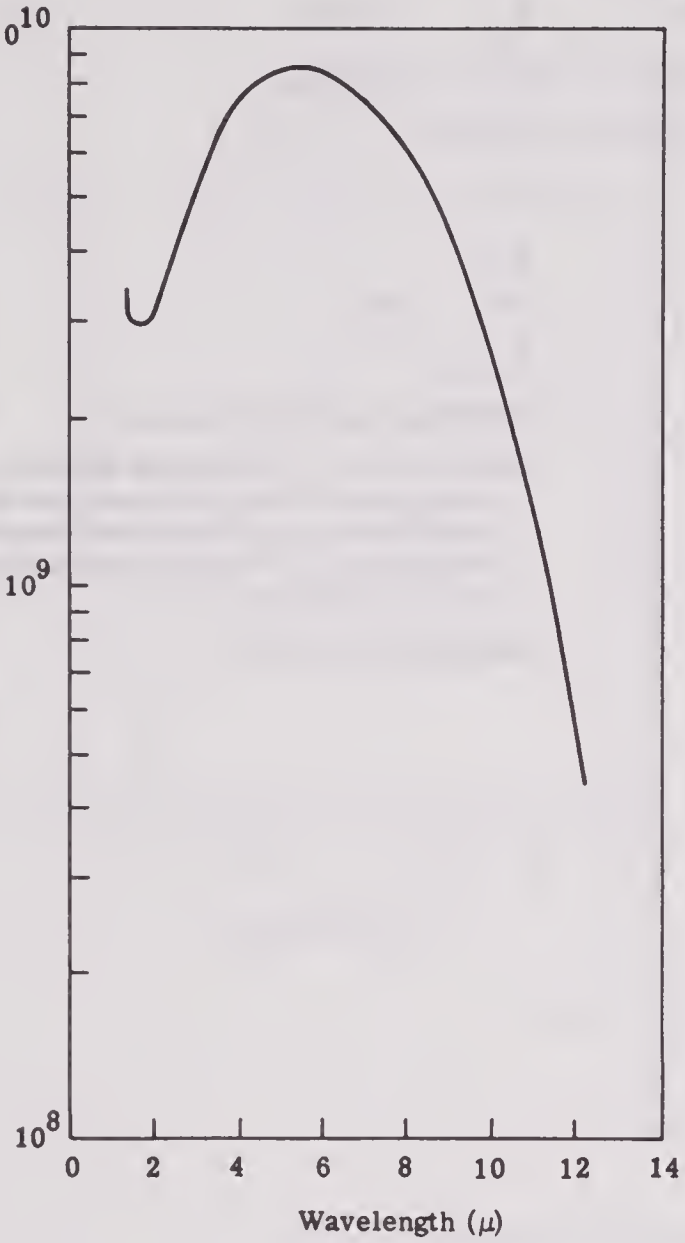
Manufacturer: RCA



DATA SHEET 25

Ge-Si:Au[1]

Detector Temperature 50°K



T_B	300°K
τ	<1 μsec
$\overline{R}$	~40 × 10^6 Ω
Z	
A	
fov	70°
$R_{\lambda}{}_{max}$	

Operating mode: photoconductive

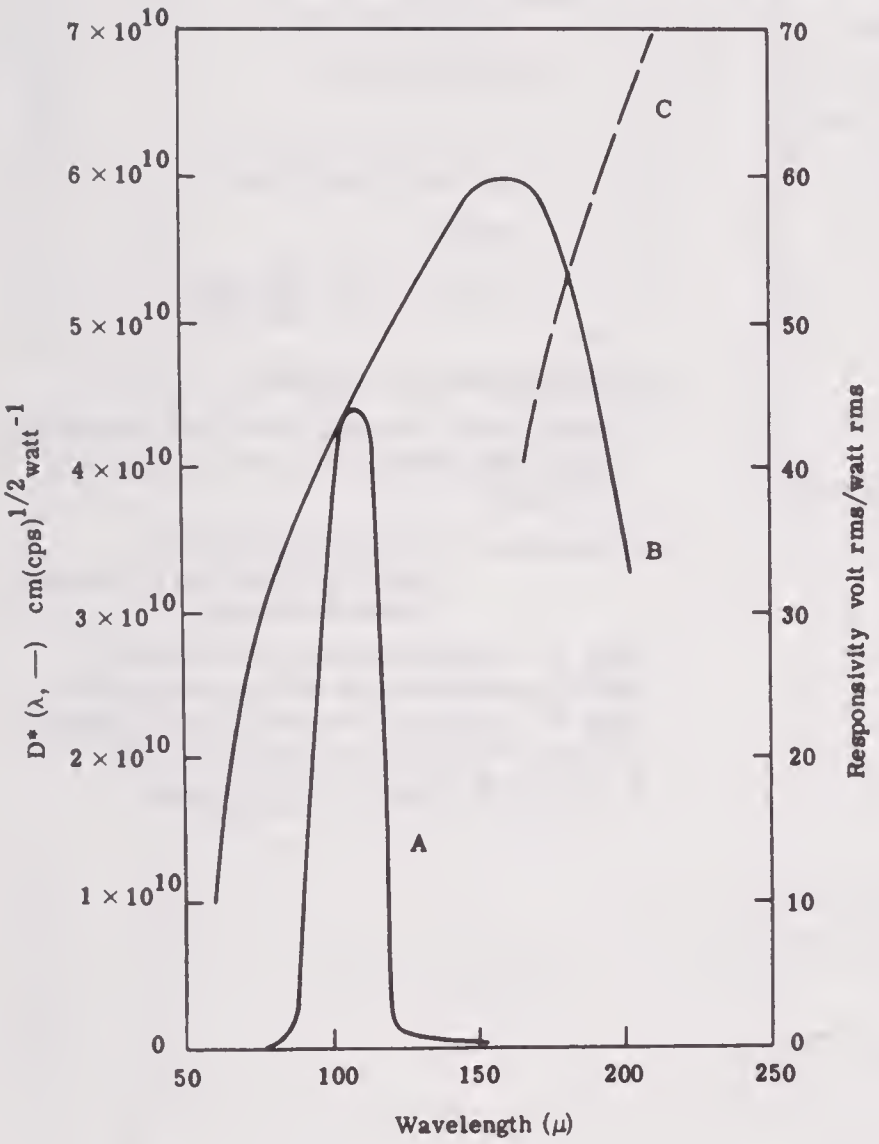
Limiting noise: presumably 1/f noise for f <100 cps, G-R noise due to photons and lattice excitations for f >100 cps.

Manufacturer: RCA

DATA SHEET 26

InSb Bolometers*

Detector Temperature (see table below)



T_B 300°K

τ <μsec

$\overline{R}$

Z

A

fov

$R_{\lambda_{max}}$

Operating mode: photoconductive

The spectral response may be tuned from 60 μ-300 μ by varying the intensity of the magnetic field

Limiting noise: unknown

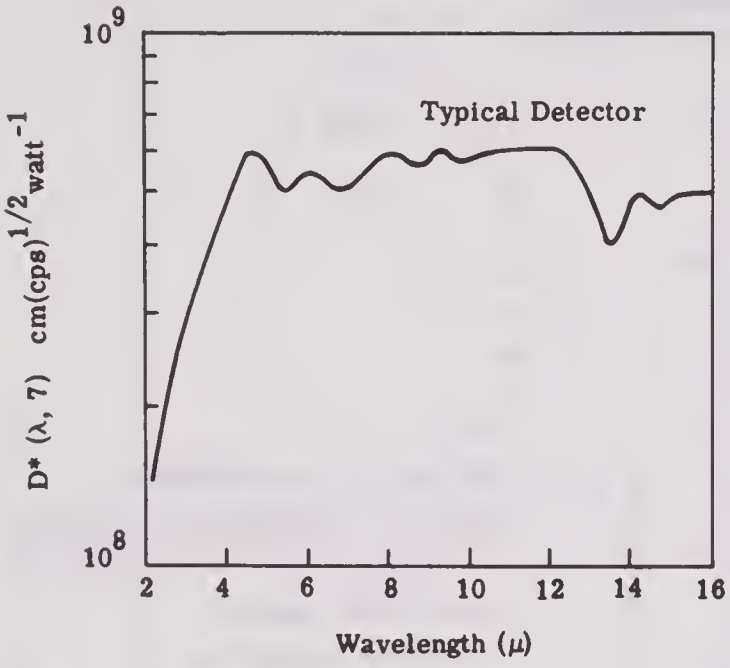
Manufacturer: Mullard Ltd.

Detector	Operating Temperature (°K)	Magnetic Field Gauss
A	4.2°K	14 K
B	4.2°K	—
C	1.8°K	5.5 K

DATA SHEET 27

Thermistors*

Detector Temperature 300°K



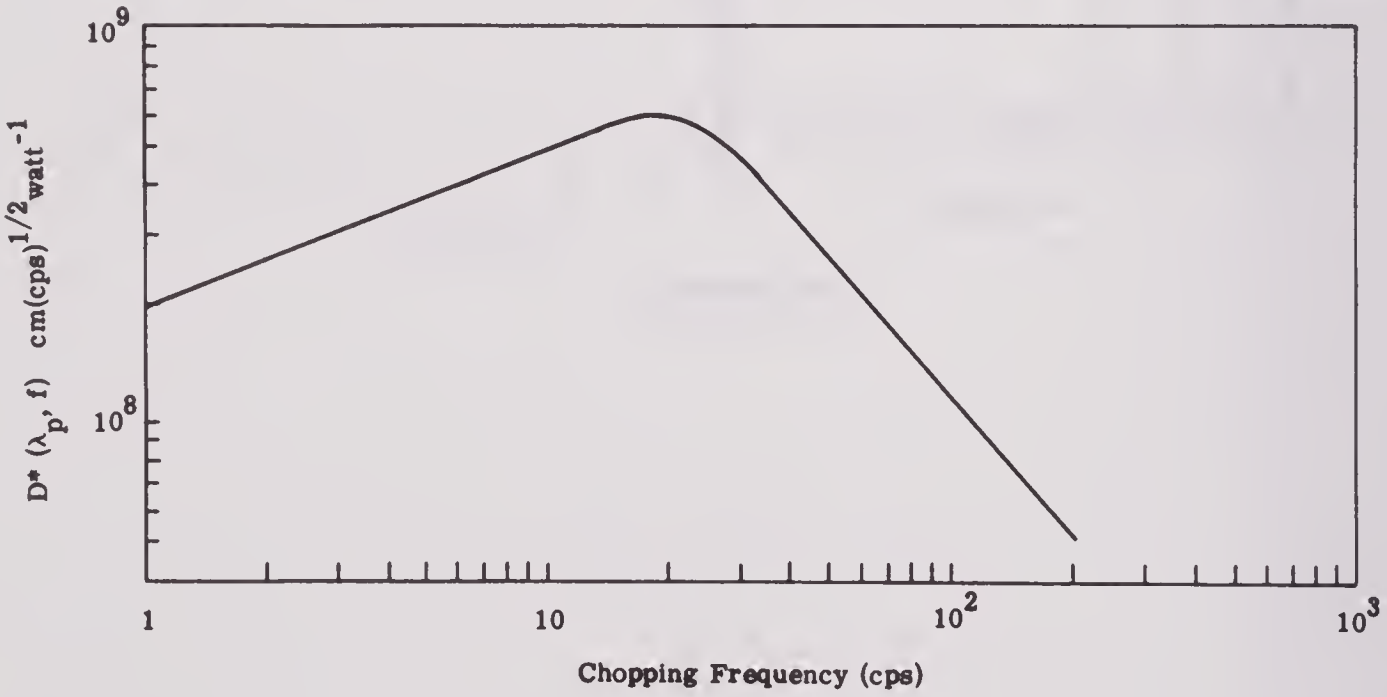
T_B	300°K
τ	800 to 8000 μsec
$\overline{R}$	10^5 to 10^6 Ω
Z	
A	10^{-4} to 3×10^{-1} cm^2
fov	180°
$R_{\lambda_{max}}$	10^2 to 5×10^3 $\frac{\text{volt rms}}{\text{watt rms}}$

Operating mode: bolometric

Limiting noise: Johnson noise for frequencies larger than about 40 cps and 1/f noise for f $\leq$40 cps.

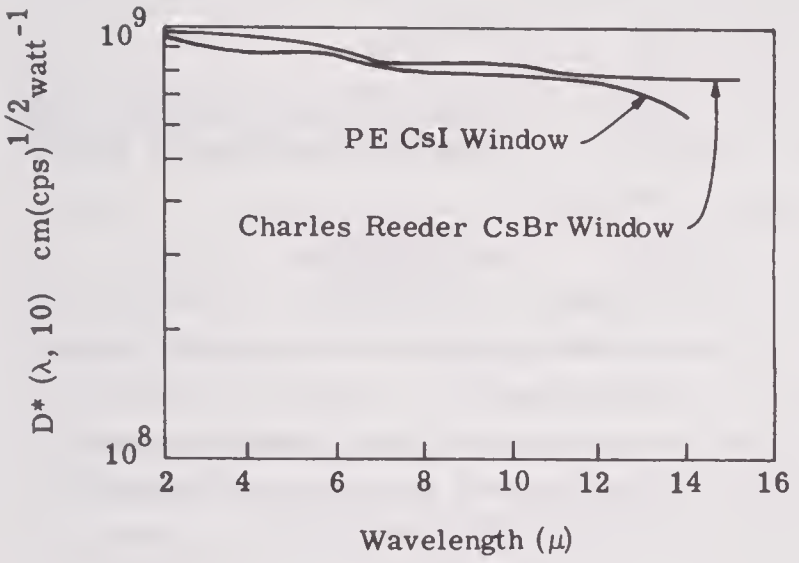
Manufacturers: Barnes Engineering
Servo Corporation of America
Polan Industries

*Note 1: immersion lenses are available which provide a factor of 3 increase in D^*.
Note 2: in general, tradeoffs can be obtained between D^* and time constant (e.g., $D^* = 8 \times 10^9 \tau^{1/2}$, where τ is in seconds).



Thermocouples

Detector Temperature 300°K

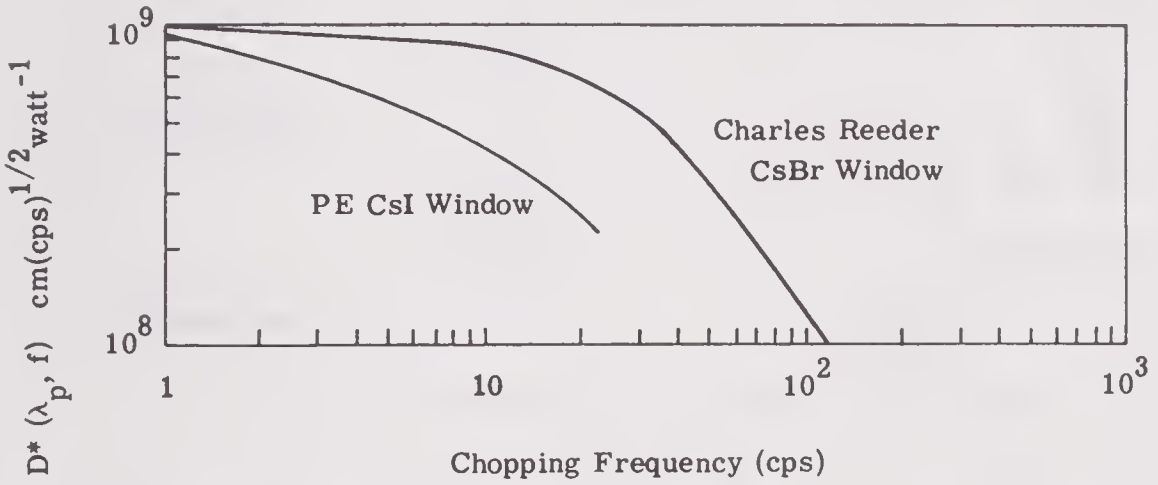


T_B	300°K
τ	~10-20 msec
$\bar{R}$	5-15 Ω
Z	
A	$(.01 \times .1)$ cm^2 to $(.03 \times .3)$ cm^2
fov	180°
$R_{\lambda_{max}}$	~5 $\frac{\text{volt rms}}{\text{watt rms}}$

Operating mode: thermovoltaic

Limiting noise: Johnson noise

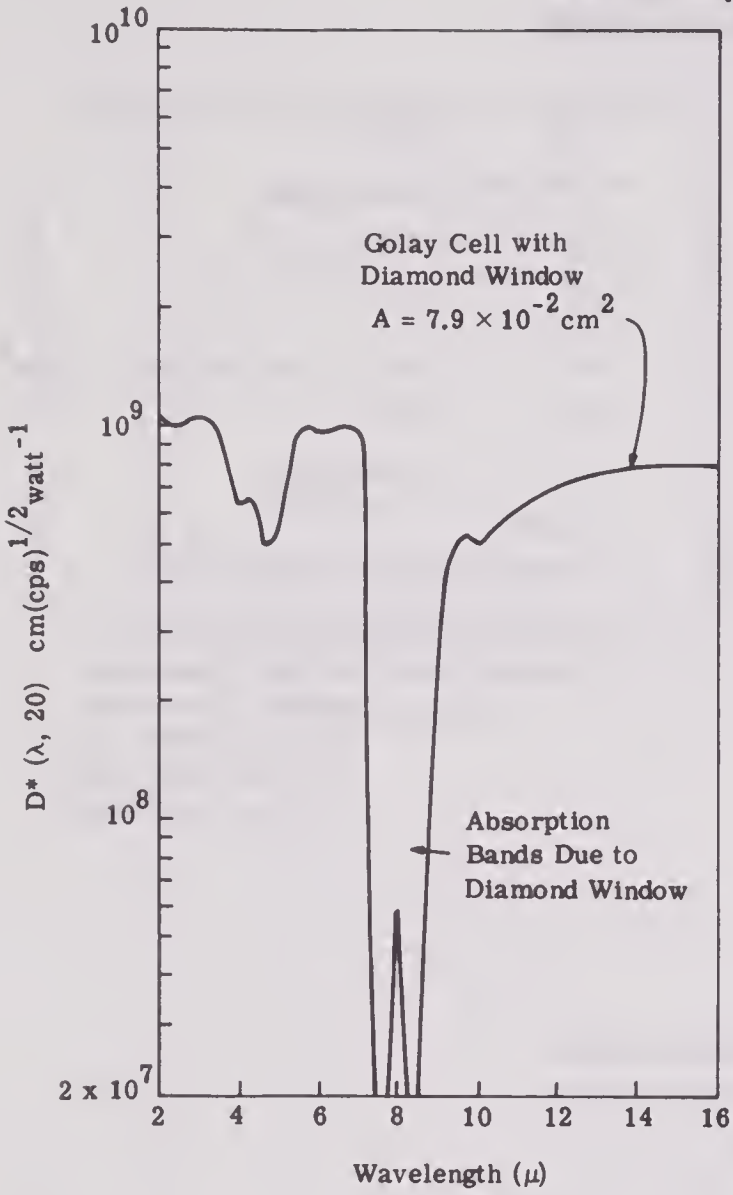
Manufacturers: Perkin Elmer Corp.
Eppley Laboratory
Charles Reeder Co.
Beckman Instruments
Farrand Corporation

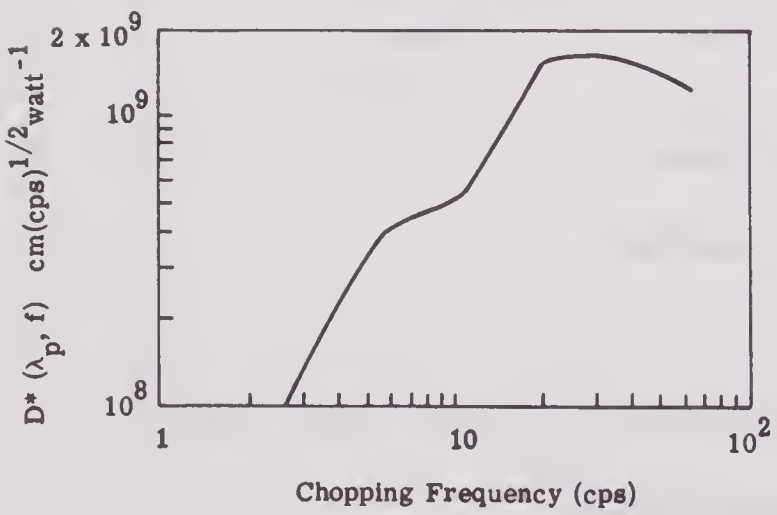


DATA SHEET 29

Golay Cell*

Detector Temperature 300°K





T_B	300°K
τ	~10 msec
$\overline{R}$	
Z, A	this device uses a pneumatic circuit coupled to a photoemissive detector.
fov	
$R_{\lambda_{max}}$	$\sim 10^4 \frac{\text{volt rms}}{\text{watt rms}}$

Operating mode: thermopneumatic

Limiting noise:

Manufacturers: Eppley Laboratory, Inc.

*Note: window materials such as diamond and quartz are available.

11.4. Test Procedures

Experimental procedures to provide the necessary information for proper detector usage are presented. Most of the experimental detail supplied here is similar to that established at the Naval Ordnance Laboratory, Corona, California [10], and Syracuse University [11]. These facilities have been sponsored by the Armed Services to provide up-to-date quantitative measurements on all types of photodetectors in the spirit of a standards laboratory with respect to the experimental procedures undertaken and the data provided. Attention should also be called to the standardization report by Jones et al. [12], which crystalizes the thinking of British, American, and Canadian scientists on the subject of testing and of describing test results.

11.4.1. Determination of NEP. The circuitry used for measurement of $V_{s,\,rms}$ and $V_{n,\,rms}$ and for the determination of optimum bias is shown in block form in Fig. 11-1. The important components of this circuitry are the infrared source (blackbody), the preamplifier, and a wideband wave analyzer. The blackbody emitter has precision temperature controls. A standard technique used in test procedures is to make measurements with a blackbody set at a temperature of 500° K. The source is mechanically modulated by a disc-type chopper. Generally, this chopper is arranged with two speeds to provide radiation modulated at 90 cps or 900 cps. This chopped radiation generates an electrical signal in the detector which is amplified and then measured with the harmonic wave analyzer. The wave analyzer is also used to determine the noise level by obtaining a reading when the detector is shielded from the chopped radiation.

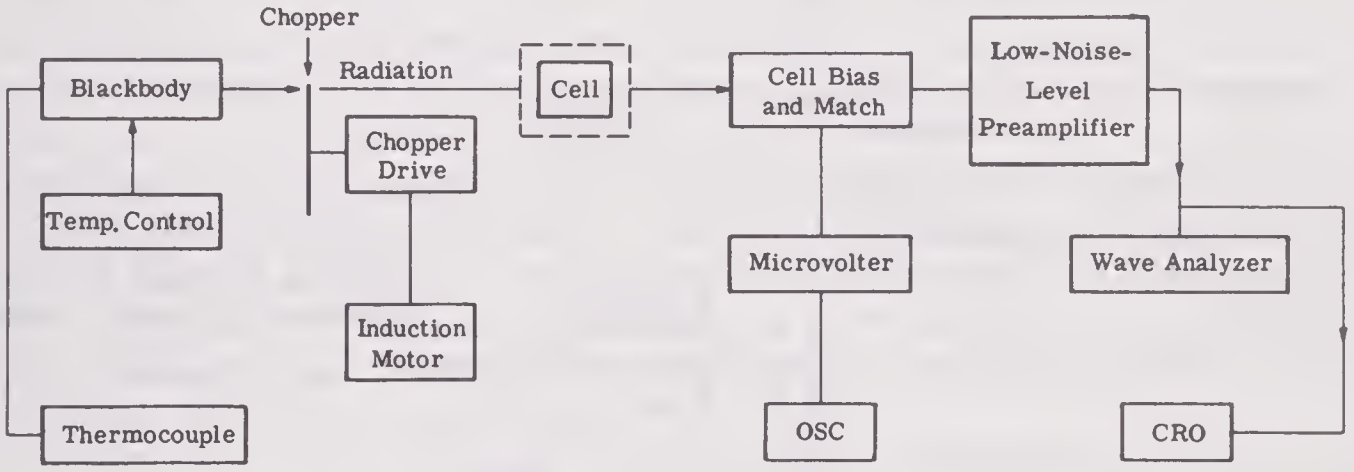


FIG. 11-1. Block diagram of test circuit used for making measurements.

The signal voltage and noise voltage for a photoconductive detector are determined as a function of bias voltage. For most detectors, chopping frequency is not a significant factor in determining optimum bias. Figure 11-2 shows a typical bias graph with plots of signal and noise voltage *versus* bias current. This graph is typical of those supplied by the Naval Ordnance Laboratory. The optimum bias point is determined from such a graph and used in all subsequent measurements for that particular detector. Typical circuitry of the cell bias and match are shown in Fig. 11-1, and drawn schematically in Fig. 11-3. Two sets of input leads are shown from the detector to the match. One set is for the bias current used with thermistors, photodetectors, and the like; the other set is for photovoltaic and PEM detectors as well as thermocouples. Photovoltaic, PEM, and thermocouple detectors usually require transformer coupling to the preamplifier because of their low impedance. It is important to use a transformer whose impedance can be varied to insure that the equivalent noise input resistance of the preamplifier is transformed to an impedance lower than the impedance of the detector being tested.

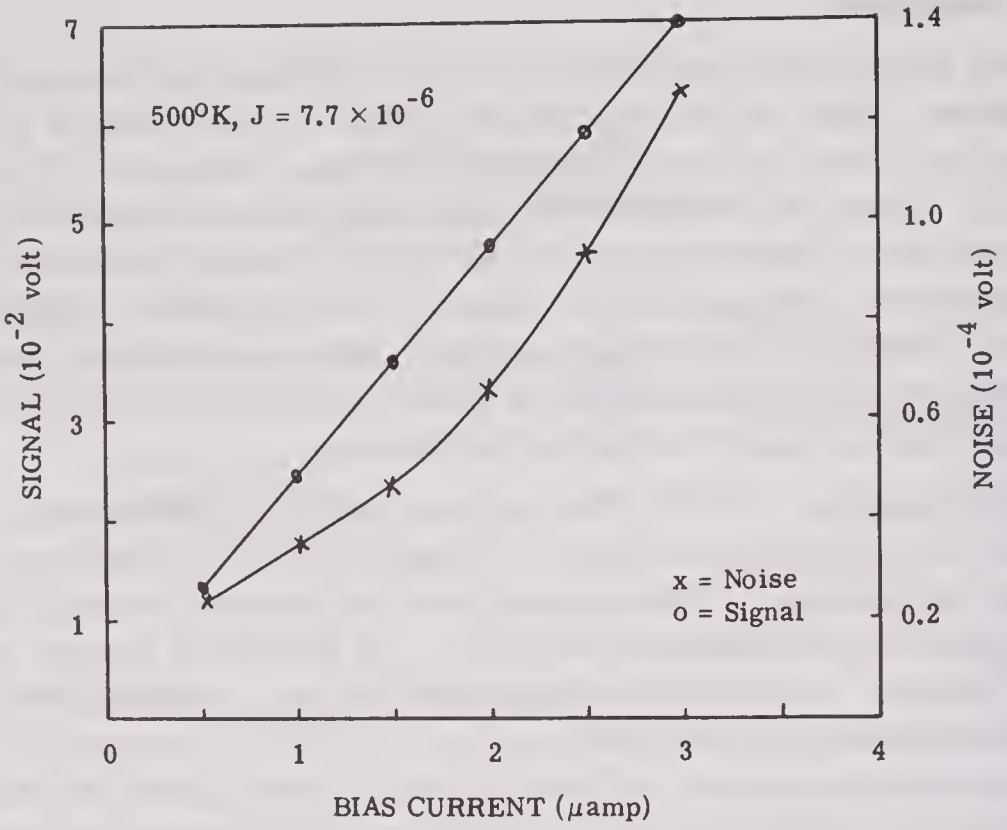


FIG. 11-2. Determination of optimum bias.

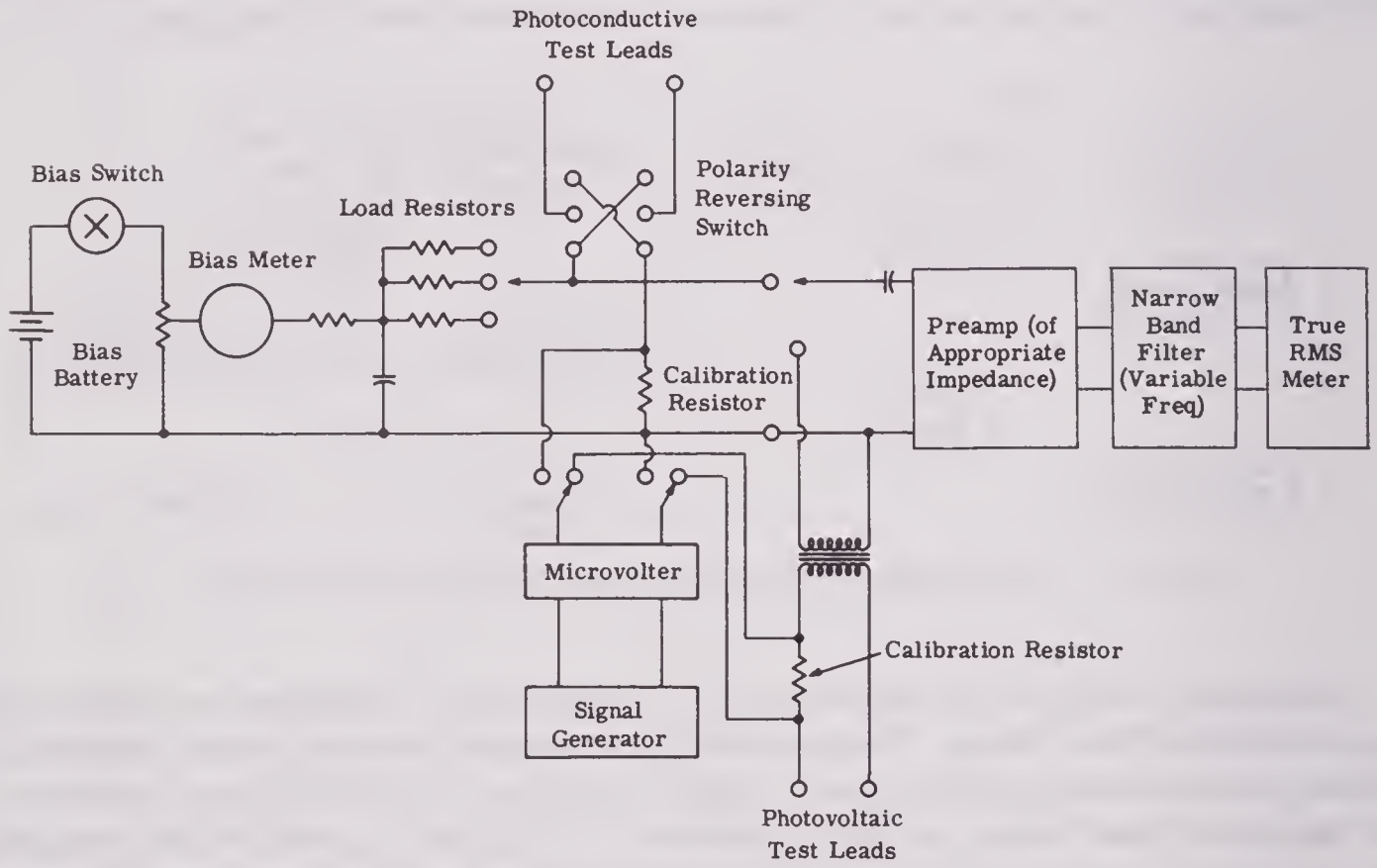
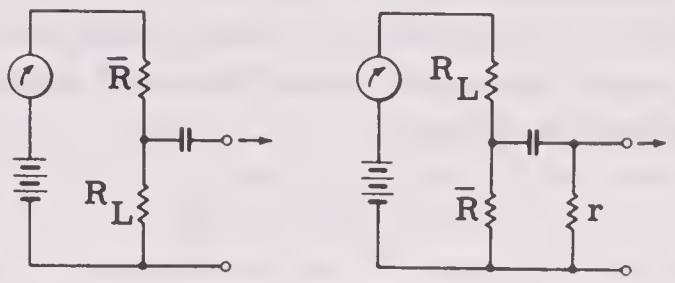


FIG. 11-3. Cell bias and impedance matching circuit for infrared detectors.

The simplest kind of circuitry associated with the photoconductive detector is shown in Fig. 11-4(*a*) and consists simply of a bias battery supply in series with the photoconductive detector and a load resistor, R_L. The signal is taken off the load resistor and fed through a capacitor to a preamplifier. The voltage across the load resistor is given by

$$V_L = V\frac{R_L}{R_L + \bar{R}} \tag{11-1}$$



(a) (b)

FIG. 11-4. Bias circuits for photoconductive detectors.

The change in voltage across the load resistor produced by the modulated radiation is found by differentiating this equation with respect to the resistance of the cell. It then follows that

$$V_L = V \frac{R_L \bar{R}}{R_L + \bar{R}} \frac{\Delta \bar{R}}{\bar{R}} \tag{11-2}$$

When the load resistance is much larger than the detector resistance, a constant bias current condition exists. Maximum signal voltage is obtained when the load resistance is the same size as the detector resistance; extended frequency response is obtained for small values of load resistance. This latter requirement usually appears with high-resistance fast detectors. In this situation capacitive effects become important, and in order to match the response-time capability of the detector it is necessary to use a small load resistor, which produces a reduced signal amplitude but flat frequency response over a wider frequency range.

A modification of this simple circuit is shown in Fig. 11-4(*b*). It involves placing a dc load resistor in series with the detector and an ac load resistor (r) across the detector through a coupling capacitor. The output is fed to the preamplifier from the ac load resistor. The effect of this type of circuitry is to allow changes in the load resistor without influencing the photoconductor bias current. These changes are necessary to determine the optimum bias current.

In making any measurements, it is important that the preamplifier noise be less than the detector noise. Two types of noise must be considered with respect to the preamplifier: (1) an effective series noise; (2) an effective shunt noise.

The series noise is experimentally determined by shorting the input to the preamplifier and then noting the output noise voltage. The shunt noise is determined by opening the input circuit and recording the output noise level. Then, starting with large values of resistance, a sequence of resistors of decreasing value is placed across the input to the preamplifier and the output noise level noted. The value of resistance is reduced to a point where a change in the noise output from the open-circuit condition of the preamplifier is recorded. One must then insure that the detector resistance is a smaller value than this to provide a detector noise greater than the preamplifier noise. In this condition the detector resistance is greater than the series resistance of the preamplifier, but less than the shunt resistance. This is referred to as a detector-noise-limited condition. This requirement becomes difficult to meet when one is forced to deal with very-high-impedance detectors (higher than 15 or 20 megohms). Otherwise, the problem of shunt noise is not serious, and one usually finds that the series noise requires the most caution.

The primary purpose in measuring signal and noise voltages with the equipment of Fig. 11-1 is to determine the photodetector's noise equivalent power (NEP). This

can now be done since the power density of the radiation from the blackbody that falls on the detector is known. This value can be calculated, starting with the Stefan-Boltzmann law. The power density H from a source of radiance N at a distance X to the detector is the detector irradiance:

$$H = N\frac{A_s}{X^2} = \frac{W}{\pi} \cdot \frac{A_s}{X^2} \tag{11-3}$$

where A_s is the source area, N is the radiant flux emitted by the source per unit area per unit solid angle, and is equal to W/π (W = radiant emittance). For a circular source aperture of diameter D_s, the power density is

$$H = \frac{W}{\pi} \cdot \frac{\pi D_s^2}{4X^2} = \frac{W}{\pi}\frac{A_s}{X^2} \tag{11-4}$$

and therefore NEP is given by

$$\text{NEP} = \frac{HA_d}{V_{s,\,rms}/V_{n,\,rms}} = \frac{W}{4}\left(\frac{D_s}{X}\right)^2 \frac{V_{n,\,rms}}{V_{s,\,rms}} A_d \tag{11-5}$$

The gain of the circuitry used to determine NEP is checked with an oscillator and a microvoltmeter connected to the input of the preamplifier. The noise bandwidth of the system is determined by measuring the Johnson noise generated in a wire-wound resistance as

$$\Delta f = \overline{e^2}/4kTr \tag{11-6}$$

where k is the Boltzmann constant, T is the absolute temperature, r is the resistance, and $\overline{e^2}$ is the mean square voltage fluctuation. The signal-to-noise ratio of a detector at a given bias current is generally independent of the load resistance. However, as shown by Eq. (11-2), the signal voltage, and correspondingly the noise voltage, are functions of the load resistor. Since different applications may require different load resistors, a listing of detector signal and noise measurements must include the value of the load resistance used in making the measurements.

11.4.2. Time Constant. Speed of response information is usually provided in one of two forms. They are: (1) a plot of response *versus* frequency from which a detector time constant can be estimated, and (2) the photodecay characteristic after removal of a photoexcitation source. Information of type (1) is generally obtained by amplitude modulation of radiation from an infrared source irradiating the detector, and varying the frequency of modulation; type (2) is obtained by observing the signal wave shape of the photodetector response to periodic pulses of light. Systems for making measurements to provide the two types of information are described below.

11.4.3. Frequency Response. Frequency response is usually measured with a metallic-disc light chopper. The disc is ringed with slits spaced symmetrically, so that the opaque region and the slit region have the same width.

The modulation frequency is given by the spinning rate of the disc multiplied by the number of slits in the disc. The higher the frequency of modulation required, the higher the spinning rate, and/or the greater the number of slits cut in the disc. Increasing the number of slits results in slits of decreasing width (for any one size disc), until eventually an optical system is required to image down the infrared source onto the slit. The radiation passing through the slits is then focused onto the detector. For low-frequency operation, sinusoidal modulation can be obtained by proper selection of the chopper opening [13].

The frequency response was determined at Corona [10] by a variable-speed chopper, giving a frequency range of 100 to 40,000 cps. Radiation from a Nernst glower is

sinusoidally modulated by the chopper, and is usually filtered by a selenium-coated germanium window. The signal from the detector is measured by putting the output of a cathode follower and a preamplifier into the y-axis input of an oscilloscope. An incandescent tungsten source is simultaneously modulated by the chopper, and activates a photomultiplier whose signal is fed into a preamplifier and a tachometer; the latter's output is proportional to frequency and is put on the x axis of the oscilloscope. The oscilloscope display is photographed as the chopper slows down from its maximum speed. Syracuse University, using a wheel cut with 1400 circular holes spinning at a rate of 10,000 rpm, obtains a maximum chopping frequency of 240,000 cps. This equipment uses a Globar as the light source and an As_2S_3 (arsenic trisulfide) lens to focus the source onto the slit. For photodetectors whose response can be described by

$$V_{s,\,rms} = \frac{V_{s,\,rms}|_{f=0}}{(1+\omega^2\tau^2)^{1/2}} \tag{11-7}$$

this high-frequency chopping rate permits an evaluation of time constants as short as 0.5 μsec. Typical frequency-response data reported from NOL, Corona, is shown in Fig. 11-5.

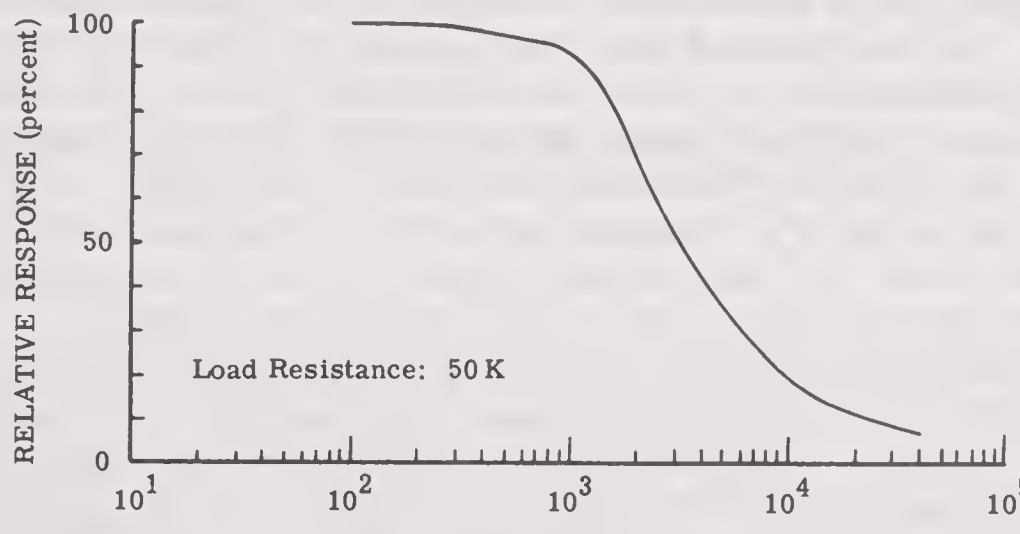


FIG. 11-5. Detector frequency response.

Frequency response measurements may also be obtained by using an injection laser as the light source. The injection current may be modulated sinusoidally at frequencies up to the gigamegacycle region—sufficient for determining time constants shorter than 1 μsec. However, the laser must be calibrated in terms of output intensity *vs.* frequency to insure that the frequency dependent of the electrical output of the detector is due to the detector and not to the source. Several manufacturers are now using this technique to determine time constant.

11.4.4. Pulse Response. Another approach to the measurement of speed of response is a direct measurement of the decay or rise characteristics of the detector. For detectors with slow response and high sensitivity, it is fairly easy to design a mechanical light chopper with sufficient speed so that the dynamic characteristics measured belong to the photodetector, and not to the chopper. However, when one is dealing with photodetectors whose response times are less than 1 μsec, normal procedures in making this measurement become difficult. To measure the decay or rise characteristics of the detector requires a light source whose rise or fall time is approximately 1/10 the time that is to be measured. Optical spinning mirror systems can provide such rapid rise-and-fall light-pulse time. A rather simple arrangement is shown in Fig. 11-6.

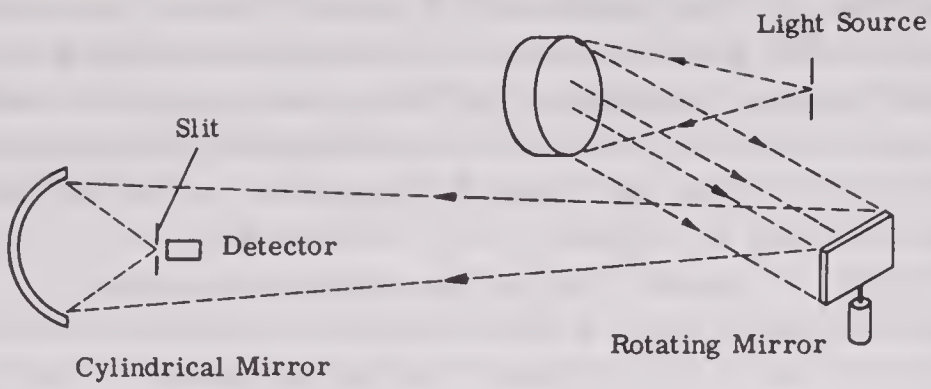


FIG. 11-6. Simple spinning mirror for periodic light pulse generation.

A collimated beam of light is deflected by a rotating mirror. A decollimating mirror which focuses the infrared radiation on the detector is placed a distance X away from the mirror. The rise time of the light pulse is the time it takes the leading edge of the pulse to fill the decollimating mirror, and the fall time is the time required for the trailing edge of the light beam to move off of that same mirror. The velocity with which the light ray moves across this mirror is given by the distance between the spinning mirror and the decollimator, multiplied by the angular velocity of the spinning mirror. The rise time and decay time, assuming a symmetrical light pulse, are equal to each other, and to the width of the decollimator divided by the velocity. Obviously, by making X sufficiently large, the rise and fall times can be made shorter, but generally at the expense of decreasing intensity at the detector. The energy may be increased by the use of a cylindrical mirror which compresses without affecting its width. Light pulses with rise and decay times of about 50 nanoseconds have been generated with this technique, using a mirror spinning at 10,000 rmp.

Another useful spinning-mirror technique is that described by Garbuny *et al.* [14]. The method consists of surrounding a rotating multisided mirror by a set of stationary mirrors (see Fig. 11-7). This assembly is so adjusted that the collimated light from the source is repeatedly reflected between the central and the stationary mirrors. Each face of the mirror rotating with angular velocity ω adds 2N to the rotational speed of the emerging light beam. If D is the width of a slit in the image plane, and is less than the width of the light beam δ, the rise time and the fall time of the pulse are given by

$$\tau_{rise} = \frac{D}{2N\omega X} \tag{11-8}$$

and

$$\tau_{decay} = \frac{\delta - aD}{2N\omega X} \tag{11-9}$$

where N is the number of faces on the rotating mirror and X its distance from the image. By using a multisided spinning mirror to obtain high tangential velocities, it is possible to substantially reduce the radial distance from the spinning mirror to the detector over that required in a construction like Fig. 11-6. Using mirror optics for collimating the light source permits any infrared emitter to be used. With a 0.5-mm wide sensitive element, a spinning-mirror rotation rate of 10,000 rmp, $X = 1$ m, and $N = 6$, rise times of 30 nanosec are readily available. Using a turbine-drive motor system to spin the mirror, rotating speeds as high as 3000 rps can be obtained, so that pulse rise and decay time of less than 1 nanosec are possible.

Square pulses of radiation may also be obtained by modulating injection lasers with square pulses of current. In the fashion described above, the rise time and decay time of the detector can then be observed.

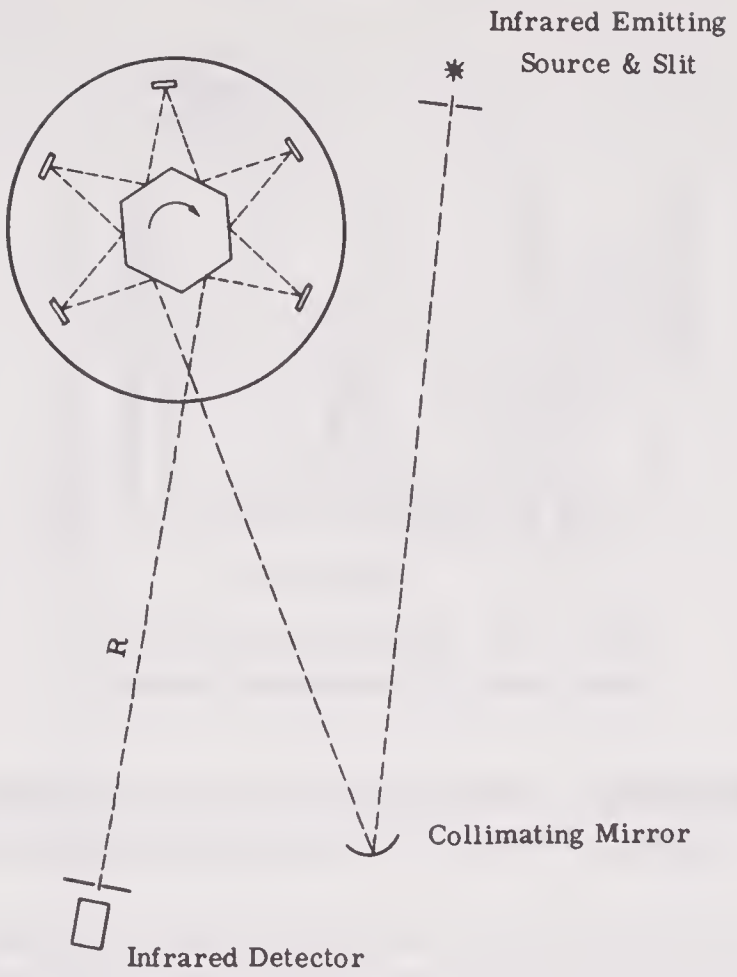


FIG. 11-7. Spinning mirror system for periodic light pulse generation in the millimicrosecond range.

Conditions exist where signal-to-noise ratios of close to or less than unity must be measured. Examples of such cases follow. (1) Examination of fractional-microsecond signals from high-impedance detectors. The measurement technique here requires the ac loading of the detector (see Fig. 11-4*b*) with a resistance low enough to provide flat frequency response over the spectrum of interest. (2) Examination of fractional-microsecond signals from low-impedance detectors such as indium antimonide. This measurement is difficult because the noise level of a wideband preamplifier is higher than that of the detector. (3) Reproduction of low-level signals caused by low-level radiation sources, as for wavelength-dependent measurements.

A device has been developed which makes measurement in these cases readily possible, and with signal-to-noise ratios less than 1. This device (often called a synchronous detector) applies a sampling technique and integration (or averaging) to the direct measurement of the shape of periodic noise-limited waveforms. This may be compared to the usual coherent detector which can be described with reference to Fig. 11-8. The signal wave shape is periodic, and is triggered in the same manner that would be required for good high-speed oscillographic reproduction; the noise is random. The interval Δt represents the "on" time of an electronic switch, during which the signal and noise voltage is fed directly into an integrator. By sampling successive intervals and averaging, it is possible to reduce the noise-voltage fluctuation observed at the integrator output without affecting the signal level. Quantitatively, the noise-voltage fluctuations are reduced by $1/\sqrt{N}$, where N is the number of observations made during the measurement. The signal-to-noise voltage ratio is then improved by the square root of N. If Δt is made small compared to the signal transient time, and is slowly and uniformly retarded in time with respect to signal onset, an accurate chart record of the signal wave shape may be produced [15].

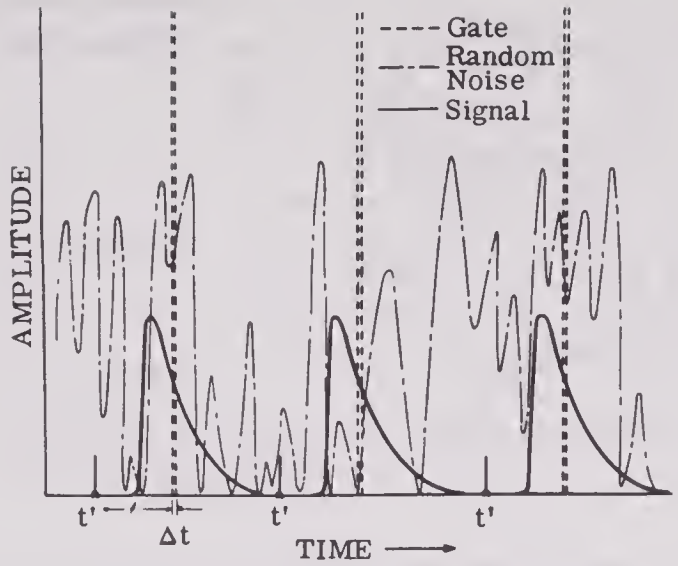


FIG. 11-8. Signal, noise, and gate relationship in waveshape recorder.

11.4.5. Spectral Response. Measurements of the wavelength dependence of infrared photodetectors are generally made with an experimental setup like the one illustrated in Fig. 11-9.

The energy flux from the exit slit is measured at each wavelength with a thermopile or thermocouple. As the wavelength is changed, the energy falling on the thermocouple is rasied or lowered to a convenient value by opening or closing the entrance slit of the monochromator, with the middle and exit slits usually remaining fixed. Once this level is set, the energy flux is allowed to fall onto the detector, and the response is then obtained. A typical relative response curve from Corona is shown in Fig. 11-10. Generally, the chopping rate of the light input to the monochromator is 10 to 13 cps, compatible with the response characteristics of the thermocouple. However, since most photodetectors show considerable improvement of NEP at higher chopping rates, it is advantageous when possible to modulate the spectral radiation at frequencies of about 200 cps. At Syracuse, the chopper is operated at 208 cps, and the detector signal is measured by feeding it through a filter of 30-cps bandwidth tuned to 208 cps to a preamplifier, and then to a vacuum-tube voltmeter. At the low chopping frequency, the detector signal is fed directly into the amplifying system of the monochromator.

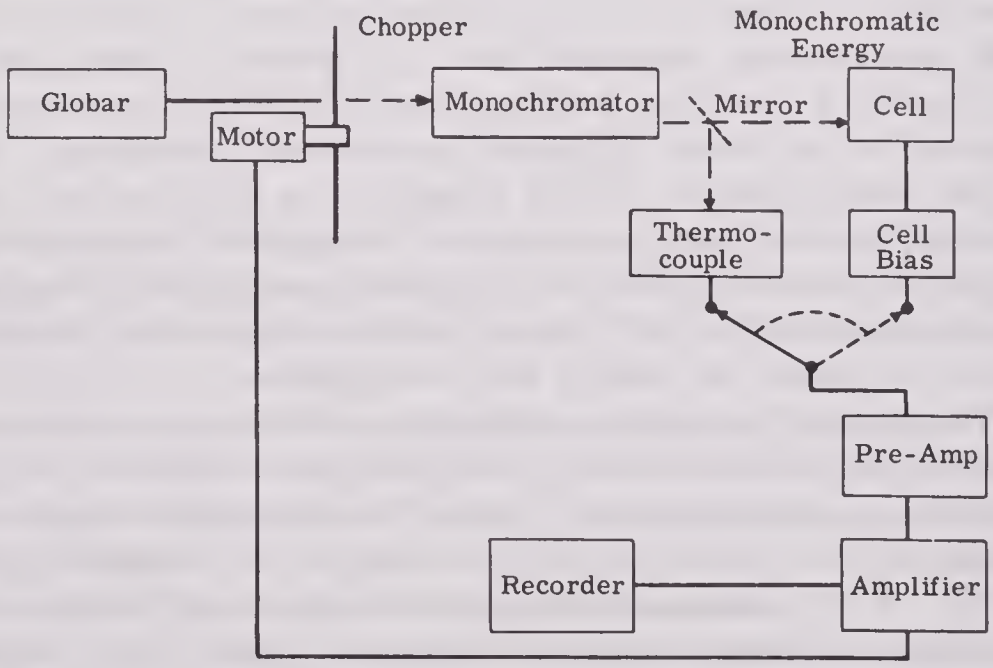


FIG. 11-9. Block diagram test circuit used for measuring detector response.

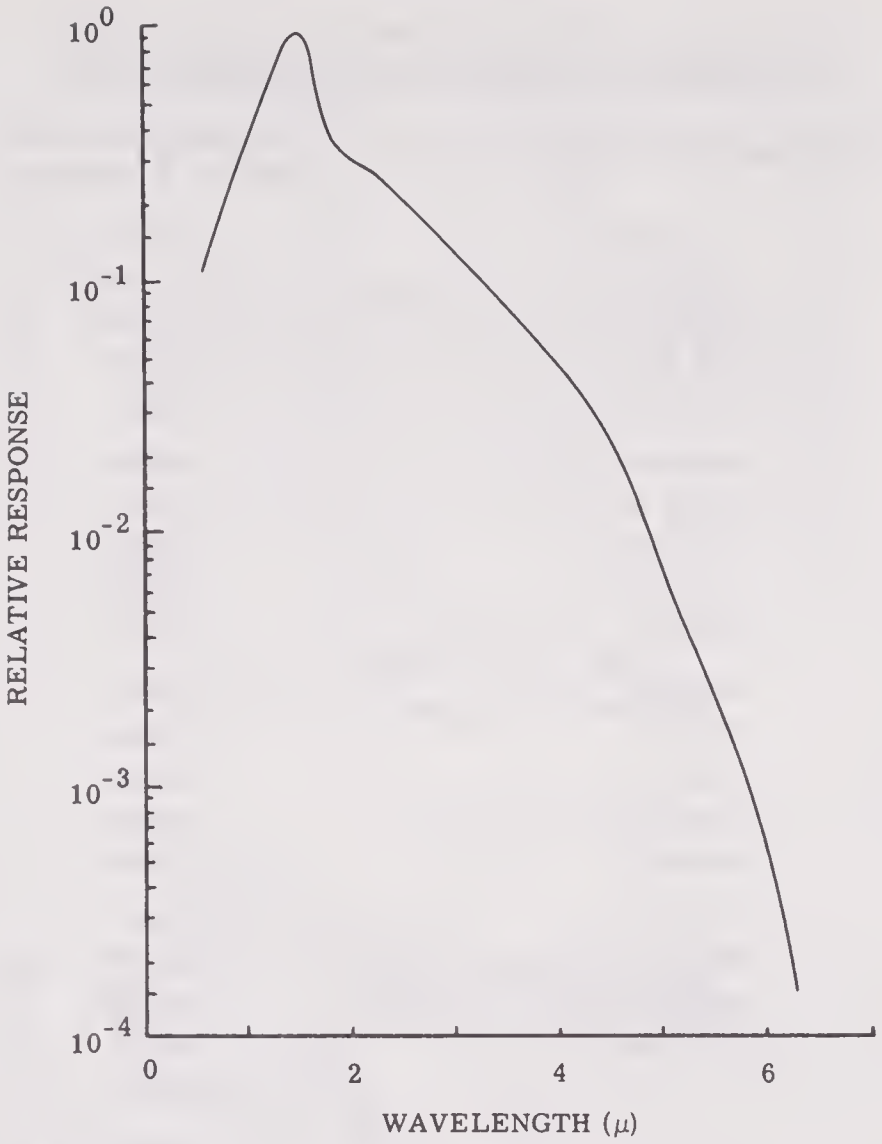


FIG. 11-10. Typical detector relative spectral response.

Along with the measured relative spectral response curve, it is important that the detector user be provided with an absolute calibration sufficiently universal that the spectral dependence of figures of merit such as NEP and D^* can be readily derived. The information available from the measurements of NEP and relative spectral response, and the theoretical law for blackbody spectral radiation distribution, are sufficient to provide the absolute calibration.

Absolute spectral measurements may be obtained using a calibrated thermocouple. However, the detector being measured must be placed at a point in the monochromatic beam, where it receives the same energy as the calibrated thermocouple. Since this is difficult to do in practice, a more suitable method is common. In this method use is made of the relationship

$$D^*(\lambda f_0) = KR_\lambda' \tag{11-10}$$

where R_λ' is the relative response and K is a proportionality constant expressed as

$$K = \frac{D^*(T_{BB}, f_0)}{\sum_{i=0}^{\infty}{}' F_{\lambda_i} R_{\lambda_i}} \tag{11-11}$$

Here, F_{λ_i} is the fraction of energy in a particular wavelength interval ($\Delta\lambda$) of the spectrum of the blackbody used for detectivity measurements and corrected for the radiation emitted by the chopper which is at 300° K. Values of F_{λ_i} for wavelengths between 1 and 30μ are given in Table 11-3. More general tables are given by Lowan and Blanch [16]. The constant K must be evaluated for each detector.

TABLE 11-3. ENERGY FRACTIONS FOR A 500°K BLACKBODY IN A 300°K SURROUNDING MEDIUM

Wavelength Interval (μ)	*Energy Fraction* (500° K *Blackbody*)
1-1.5	7×10^{-6}
1.5-2.0	3.7×10^{-4}
2.0-2.5	0.0032
2.5-3.0	0.012
3.0-3.5	0.024
3.5-4.0	0.038
4.0-4.5	0.050
4.5-5.0	0.058
5.0-5.5	0.062
5.5-6.0	0.063
6.0-6.5	0.061
6.5-7.0	0.058
7.0-7.5	0.054
7.5-8.0	0.050
8.0-8.5	0.045
8.5-9.0	0.041
9.0-9.5	0.037
9.5-10.0	0.033
10.0-10.5	0.029
10.5-11.0	0.027
11-12	0.045
12-13	0.035
13-14	0.029
14-15	0.022
15-16	0.019
16-17	0.015
17-18	0.013
18-19	0.011
19-20	0.0084
20-22	0.015
22-24	0.0097
24-26	0.0072
26-28	0.0058
28-30	0.0029

11.4.6. Noise Spectrum. The noise-voltage spectrum is obtained with the system described in Section 11.3.1. However, the light source is removed and the noise voltage is obtained by simply reading the voltage at the wave analyzer. A typical plot of noise spectrum is shown in Fig. 11-11.

11.4.7. Sensitivity Contours. If a microscopic ray of light is projected onto the surface of a photodetector, and the photoresponse recorded as a function of the ray's position, it is found that the photoresponse generally changes with the ray's position. The surface of the detector is thus rarely uniform in its photoresponse. The film detectors (lead compound family) are the worse offenders in this regard. If a graph of photoresponse *versus* light-ray position is made, and points of equal photoresponse are linked together, the resultant plot provides a "sensitivity contour," illustrated by Fig. 11-12.

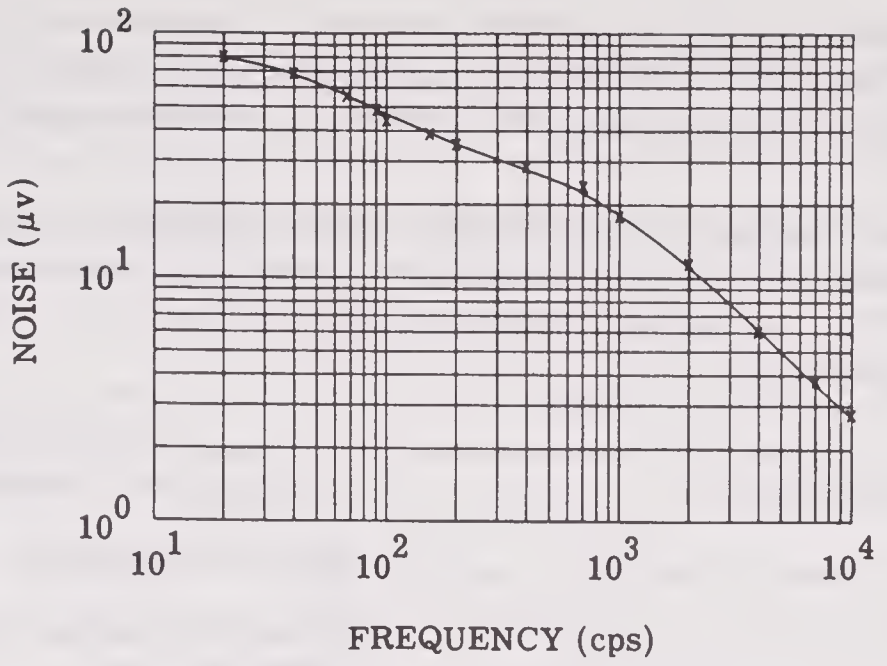


FIG. 11-11. Detector noise spectrum.

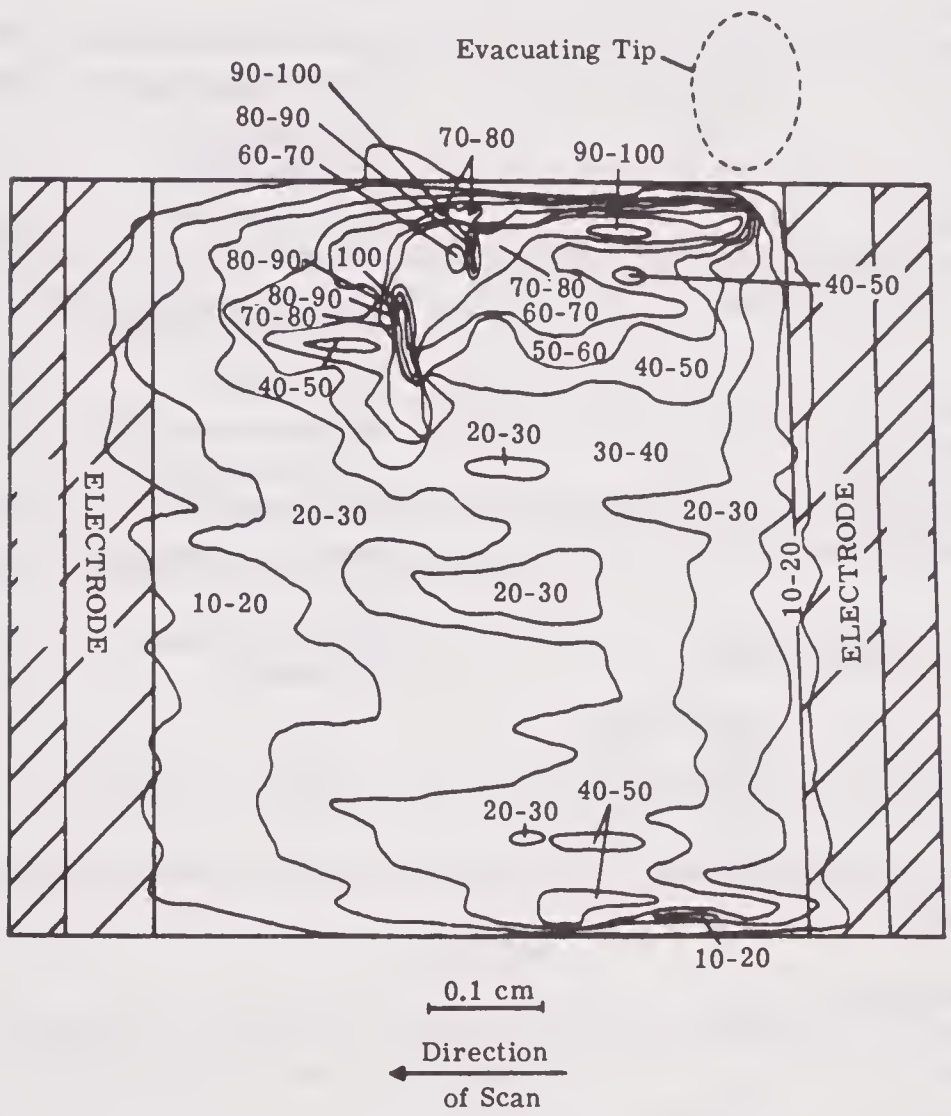


FIG. 11-12. Sensitivity contour for typical *PbSe* cell.

At Corona, the experimental arrangement to obtain sensitivity contours uses a microtable which allows the cell to be moved a small measured amount. The table is linked through a system of gears to a plotting table which gives an increase in the scale up to 36:1. The exciting radiation is from an incandescent tungsten bulb chopped at 90 cps; it is passed in reverse through a microscope such that a spot 0.066 mm in diameter is focused onto the detector. As the detector is moved beneath this radiation, the relative response at 10% intervals is noted on the plotting table. Lines connecting

equal points of sensitivity are then drawn to obtain a plot such as Fig. 11-12. This light-probe technique is also important for its utility in fundamental research programs on detector materials, where it is used in studies of diffusion length, time constant, and mobility [17].

11.4.8. General Comments. A summation of data necessary to evaluate a detector is shown in Fig. 11-13, which consists of a typical data sheet from an NOLC report. The definitions of the various parameters are listed in Table 11-1. Data on specific detectors is in Section 11.2.

TEST RESULTS		CONDITIONS OF MEASUREMENT	
R (volts, watt) (500, 90)	4.1×10^4	Blackbody temperature (^oK)	500
H_N (watts/cps$^{1/2}$ · cm^2) (500, 90)	8.8×10^{-9}	Blackbody flux density (μwatts/cm^2, rms)	9.0
P_N (watts/cps$^{1/2}$) (500, 90)	5.6×10^{-11}	Chopping frequency (cps)	90
		Noise bandwidth (cps)	5
D* (cm-cps$^{1/2}$) (500, 90)	1.4×10^9	Cell temperature (^oK)	197
Responsive time constant (μsec)	26	Cell current for 90-cps data (μa)	7.0
$\frac{R_{\lambda\ max}}{R_{bb}}$	9.1	Cell current for D^*_{mm}(μa)	20.0
		Load resistance (ohms)	2.5×10^6
Peak wavelength (μ)	2.2	Transformer	---
Peak detective modulation frequency (cps)	4×10^3	Relative humidity (%)	16
		Responsive plane (from window)	---
D^*_{mm}(cm cps $^{1/2}$/watt)	2.8×10^{10}	Ambient temperature (^oC)	24
CELL DESCRIPTION		Ambient radiation on detector	297^{o}K only
Type	PbSe (evap.)		
Shape of sensitive area (cm)	0.038×0.168		
Area (cm^2)	6.3×10^{-3}		
Dark resistance (ohms)	1.64×10^6		
Dynamic resistance (ohms)	---		
Field of view	---		
Window material	Sapphire		

FIG. 11-13. Detector data sheet (*from NOLC Report* 564).

11.5. Theoretical Limit of Detectivity

The optimal performance of an infrared detector would occur when the inherent detector noise was negligible compared to the noise induced in the detector by the random arrival rate of photons coming from the target. Thus far, to the author's knowledge, this ultimate performance has not been realized. There are, however, a number of commercially available infrared detectors which under the proper operating conditions have a limiting noise due to the random arrival rate of photons from the

background. The background is composed of the atmosphere, spectral filter, window material, mirrors, and other objects besides the target in the field of view. Detectors which can give this type of performance include PbS, PbSe, InAs, InSb, Ge:Hg, Ge:Cu, Ge:Cd, Ge-Si:Au, and Ge-Si:Zn.

11.5.1. Derivation of D^* for Photon Noise Limitation. According to Planck's law, the power radiated into a hemisphere per unit wavelength from a blackbody is given by

$$W_\lambda = (2\pi c^2 h)\lambda^{-5}(e^{hc/\lambda kT} - 1)^{-1} \tag{11-12}$$

Since there are λ/hc photons $\sec^{-1}$ w^{-1} at wavelength λ, the number of photons $\sec^{-1}$ cm^{-2} μ^{-1} can be expressed as

$$n(\lambda) = (2\pi c)\lambda^{-4}(e^{hc/\lambda kT} - 1)^{-1} \tag{11-13}$$

where $n(\lambda)$ can be thought of as the average number of photons. According to the Bose-Einstein relation, the mean square fluctuation in the number of photons is given by

$$(\overline{\Delta n^2})_\lambda = \bar{n}\left(\frac{e^{hc/\lambda kT}}{e^{hc/\lambda kT} - 1}\right) \tag{11-14}$$

where $\bar{n}$ is the average number of photons. Therefore the mean square fluctuation (in photons) of wavelength λ emitted by a blackbody becomes

$$(\overline{\Delta n^2})_\lambda = \frac{2\pi c}{\lambda^4}\,\frac{1}{(e^{hc/\lambda kT} - 1)}\,\frac{e^{hc/\lambda kT}}{(e^{hc/\lambda kT} - 1)}$$

Since the detectors mentioned above are sensitive to wavelengths shorter than 30 μ, $hc/\lambda >> kT$, which means that $\exp(hc/\lambda kT) >> 1$, and this permits the following approximation:

$$(\overline{\Delta n^2})_\lambda \approx \frac{2\pi c}{\lambda^4}\, e^{-hc/\lambda kT} \tag{11-15}$$

To find the total mean square photons in the spectral region to which the detector is sensitive, we must integrate the above expression:

$$\overline{\Delta n^2} \approx 2\pi c \int_0^{\lambda_c} \lambda^{-4}\, e^{-hc/\lambda kT}\, d\lambda \tag{11-16}$$

Therefore

$$\overline{\Delta n^2} \approx 2\pi c \left(\frac{kT}{hc}\right) e^{-hc/\lambda_c kT}\left[2\left(\frac{kT}{hc}\right)^2 + \frac{2kT}{hc\lambda_c} + \frac{1}{{\lambda_c}^2}\right]$$

The rms fluctuations in the electrical bandwidth Δf and a detector area A is given by

$$\sqrt{2\,\Delta f\,\overline{\Delta n^2}\,A\eta}$$

where η is the quantum efficiency of the responsive element. The generation rate (G_s) of carriers due to signal photons is given by

$$G_s = \frac{\lambda}{hc}\eta P_s$$

where P_s is the incident signal power in rms watts. For a signal-to-noise ratio of unity, G_s must be the same as the generation rate due to the noise source, or

$$\frac{\lambda\eta P_s}{hc} = \sqrt{2\,\Delta f\,\overline{\Delta n^2}\,A\eta}$$

P_s is therefore the signal power necessary to produce a signal-to-noise ratio of unity, and is by definition the NEP.

$$\text{NEP} = P_s = \frac{hc}{\lambda\eta}\sqrt{2\,\Delta f\,\overline{\Delta n^2}\,A\eta}$$

Substituting in the expression for $\overline{\Delta n^2}$ one obtains the theoretical limit of NEP.

$$\text{NEP} = 2\sqrt{\frac{A\pi kT\,\Delta f}{\eta h}}\, e^{-hc/2\lambda_c kT}\left[2\frac{kT^2}{hc}+\frac{2kT}{hc\lambda_c}+\frac{1}{{\lambda_c}^2}\right]^{\frac{1}{2}} \tag{11-17}$$

Using the definition of D^* from Table 11-1 and the above expression:

$$D^* \approx \frac{1}{2}\sqrt{\frac{h\eta}{\pi kT}}\, e^{hc/2\lambda_c kT}\left[2\left(\frac{kT}{hc}\right)^2+\frac{2kT}{\lambda_c hc}+\frac{1}{\lambda_c}\right]^{-1/2} \tag{11-18}$$

For this derivation a detector with a viewing solid angle of π steradians is assumed. If the viewed solid angle, Ω, has circular symmetry, it can be expressed in terms of the cone angle as follows

$$\Omega = \pi \sin^2(\theta/2) \tag{11-19}$$

where θ is the full cone angle. The dependence of D^* upon this cone angle becomes

$$D^* \approx \frac{1}{2}\sqrt{\frac{\eta h}{\pi kT}}\, e^{hc/2\lambda_c kT}\left[2\left(\frac{kT\lambda_c}{hc}\right)^2+\frac{2kT\lambda_c}{hc}+1\right]^{-1/2}\lambda_c\left(\sin^{-1}\frac{\theta}{2}\right) \tag{11-20}$$

From Eq. (11-20) one can observe the effects of field of view or changes in λ_c upon D^*. The effect of changing the field of view can be seen in Fig. 11-14. Here the relative value of D^* is plotted as a function of the cone angle θ. Figure 11-15 gives D^* at λ_p as a function of long-wavelength cutoff with background temperature as a parameter. Equations (11-19) and (11-20) and Fig. 11-14 and 11-15 as they are apply to photoemissive and photovoltaic detectors. For photoconductive detectors, which are sensitive to the population of carriers in the conduction band and for which the fluctuation in recombination rates is significant, Eq. (11-20) must be divided by $\sqrt{2}$ and the necessary scaling adjustment must be made on the figures.

Similar calculations can be made for thermal detectors which are limited by the fluctuations in the absorbed power. However, in this case one must also consider the

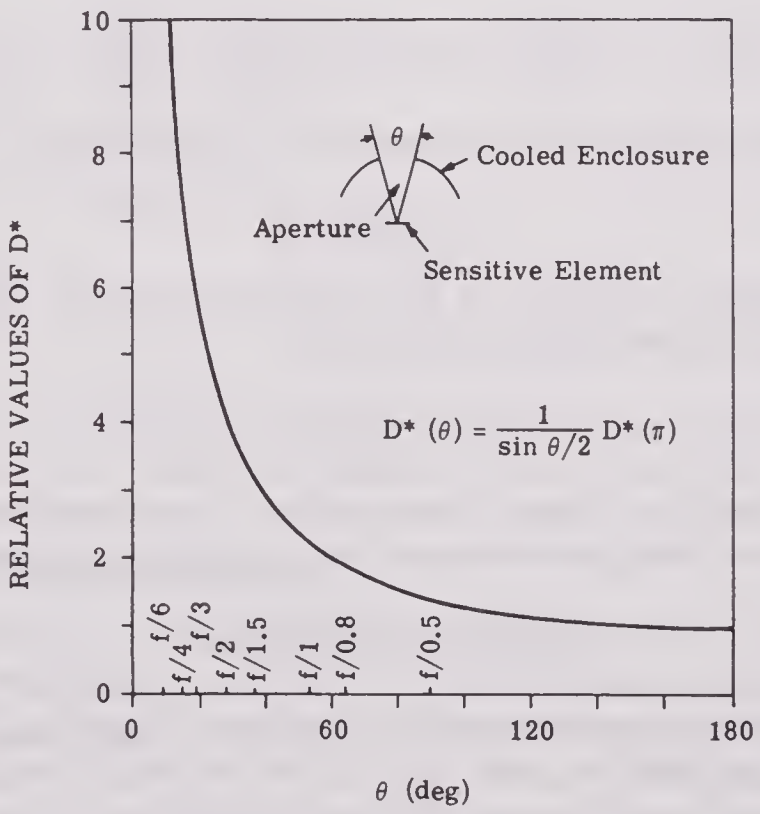


FIG. 11-14. D^* and D_a^* as a function of angular field of view.

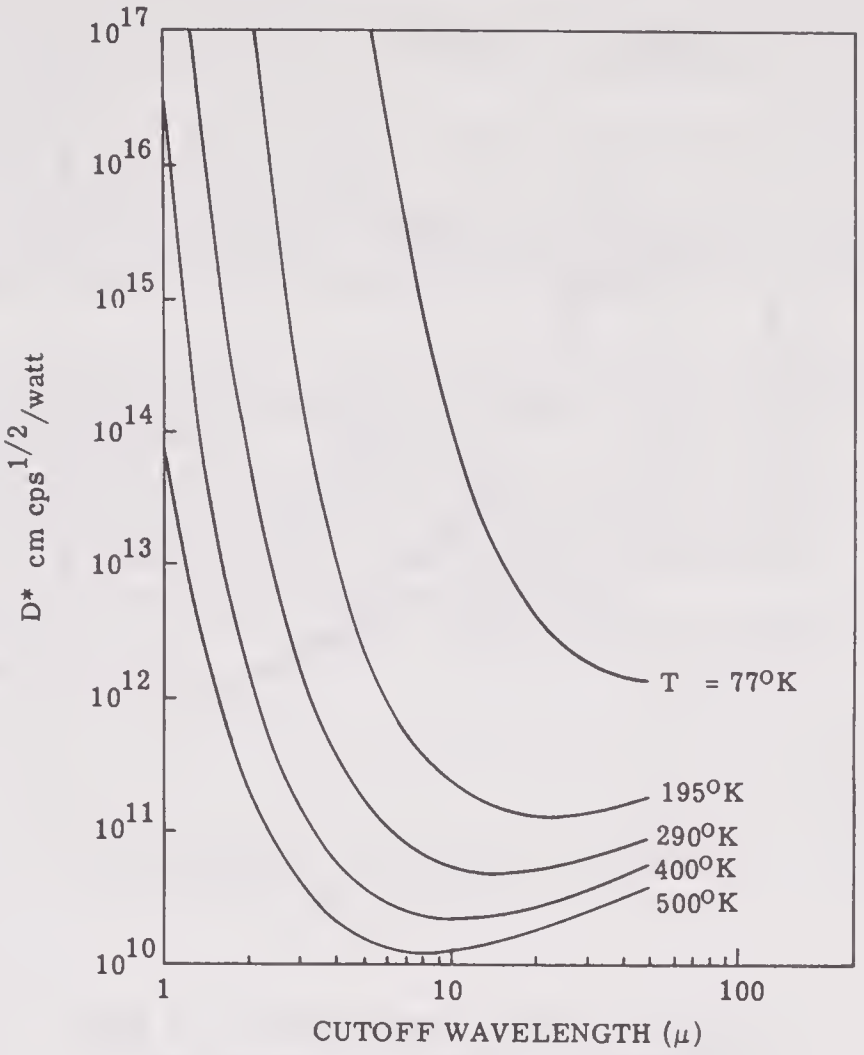


FIG. 11-15. D_a^* *versus* long wavelength cutoff for background limited detection.

contribution to the noise made by random fluctuations in the power emitted by the detector. The resulting equation for D^* is

$$D^* = \frac{4 \times 10^{16} \epsilon^{1/2}}{(T_1^5 + T_2^5)^{1/2}} \text{ cm (cps)}^{1/2} \text{ w}^{-1} \tag{11-21}$$

where T_2 is the background temperature, T_1 is the detector temperature, and ϵ is the emissivity of the sensitive element. Figure 11-16 presents D^* for photon-noise-limited thermal detectors with detector temperature as a parameter. These results assume a 180° field of view and an infinite spectral response. When cooled spectral filters are used to limit the radiation striking the sensitive element to a narrow spectral bandwidth, the theoretical limit of D^* approaches that of a photodetector sensitive to the same narrow spectral region.

11.5.2. System Design Considerations. When detectors are limited by the random arrival rate of background photons, a number of interesting problems arise. Equation (11-20) indicates that decreasing the detector field of view (reducing θ) and decreasing the background temperature T, will lead to considerable enhancements in D^*. In using this equation one must keep in mind that detector noise output contains (1) generation-recombination, (2) Johnson, and (3) $1/f$ noises. Also, one must recall that the detector resistance (for detectors which utilize changes in conductivity) is related to the number of background photons by $\overline{R} \sim 1/n$. One can see then that reducing θ and T will reduce the photon noise, but $1/f$ and the lattice contribution to generation-recombination noise will be unaffected. However, Johnson noise (J.N.) will increase in the following way:

$$V_{J.N.} \propto \sqrt{4kT\,\Delta f} \cdot \sqrt{\frac{1}{n}}$$

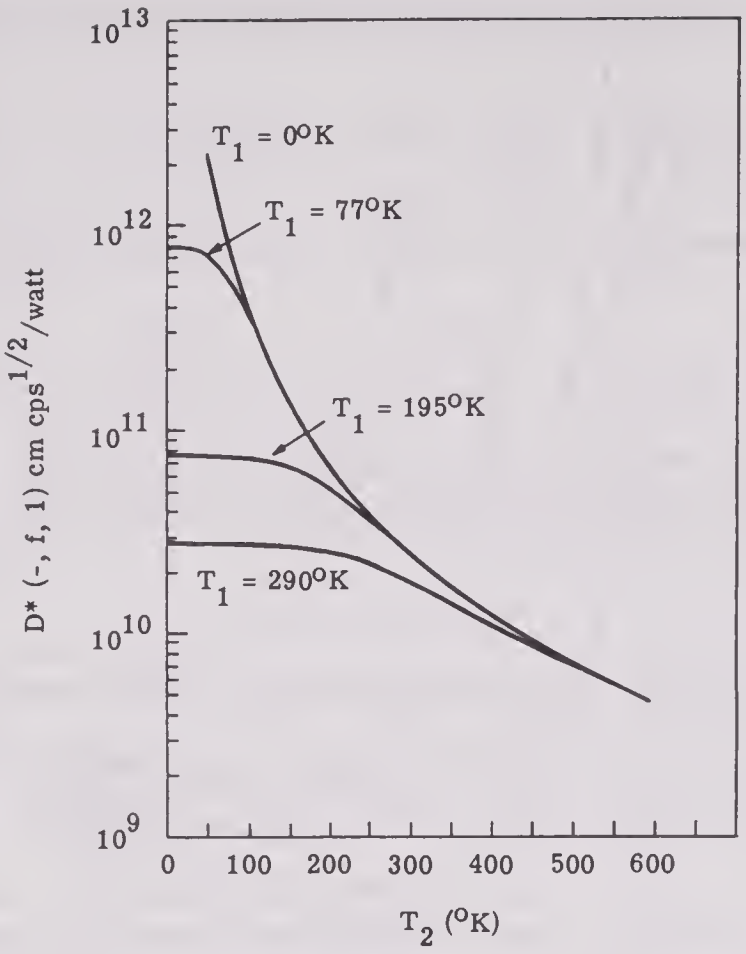


FIG. 11-16. Photon-noise-limited D^* of thermal detectors as a function of detector temperature T_1 and background temperature T_2.

One finds then that reducing θ to enhance D^* is possible within limits set by the other noise sources.

Three other effects occur which hamper the implementation of this enhanced detector performance. They are: an increased time constant due to reduction in background photon rate, the unwieldy, high value of $\overline{R}$, and the resulting low noise from the detector. The first problem sets a limit to scanning and tracking modes. The high resistance requires special care in designing the bias circuit and also in selecting the preamplifier. The reduced detector noise output places a further requirement on the preamplifier; *i.e.,* it must have a lower noise level than the detector to insure detector noise limited performance—a necessary condition to realize the full enhancements in D^*. It may be necessary to cool the preamplifier to achieve this.

In view of the above discussion, it is apparent that D^* may be improved by adjusting certain parameters, but realizing this improvement is by no means a simple task.

References

1. R. A. Smith, F. E. Jones, and R. P. Chasmar, *The Detection and Measurement of Infrared Radiation,* Clarendon Press, Oxford (1957).
2. M. Holter, et al., *Fundamentals of Infrared Technology,* Macmillan, N.Y. (1962).
3. P. W. Kruse, L. D. McGlauchlin, and R. B. McQuiston, *Elements of Infrared Technology,* Wiley N.Y. (1962).
4. W. Wolfe and T. Limperis, *IRIA State of the Art Report: Infrared Quantum Detectors,* 2389-50-T, The University of Michigan, Ann Arbor, Michigan (July 1961).
5. H. Levinstein, *Interim Report on Infrared Detectors,* **105-1,** Syracuse Research Institute (Feb. 1963).
6. *Properties of Photoconductive Detectors,* NOLC, Continuing series begun June 30, 1952.
7. *Interim Report on Infrared Detectors,* Syracuse University, Continuing series begun Jan. 1, 1954.

8. T. Limperis, "Report of the Detector Evaluation and Information Committee of the IRIS Specialty Group on Detectors," *Proc. IRIS,* Vol. IX, No. 1, January 1964, p. 165 (Unclassified paper in Secret volume).
9. *General Specifications for Environmental Testing and Associated Equipment,* U.S. Army Specification MIL-E-5272, 13 April 1959.
10. R. F. Potter, J. M. Pernett, and A. B. Naugle, *Proc. IRE* **47,** 1503 (1959).
11. P. Bratt, W. Engeler, H. Levenstein, A. MacRae, and J. Pehek, *Final Report on Ge and InSb Infrared Detectors,* Syracuse University under WADD Contract No. AF33(616)-3859 (Feb. 1960).
12. R. C. Jones, D. Goodwin, and G. Pullan, *Standard Procedure for Testing Infrared Detectors and for Describing Their Performance,* Office of Director of Defense Research and Engineering, Washington, D.C. (Sept. 1960).
13. R. B. McQuistan, *J. Opt. Soc. Am.* **48,** 63 (1958).
14. M. Garbuny, T. P. Vogl, and J. R. Hansen, *Rev. Sci. Inst.* **28,** 826 (1957).
15. S. Nudelman and J. T. Hickmott, *Bull. Am. Phys. Soc.* **4,** 153 (1959).
16. A. N. Lowan and G. Blanch, *J. Opt. Soc. Am.* **30** (1940).
17. K. Lark-Horovitz, V. A. Johnson, and L. Marten (eds.) *Methods of Experimental Physics,* Academic Press, N.Y. (1959).

Chapter 12

DETECTOR COOLING SYSTEMS*

CONTENTS

12.1. General 520
12.1.1. Limitations of Fluid Cooling 520
12.1.2. Cooling System Design Criteria. 522
12.2. Types of Detector Cooling Systems 522
12.2.1. Direct-Contact Coolers 525
12.2.2. Joule-Thomson Coolers. 532
12.2.3. Expansion-Engine Coolers 540
12.2.4. Thermoelectric Coolers. 544
12.2.5. Dewars 557
12.3. Space-Environment Cooling Systems 561
12.3.1. Operating Principles 561
12.3.2. Design 562
12.3.3. Practical Cooling Systems 563
12.4. List of Manufacturers 566

*Material prepared by the technical writing staff of McGraw-Hill, Inc.

12. Detector Cooling Systems

12.1. General

The proper operating temperature for an infrared detector is determined by studying the effect of temperature on the detector parameters, and selecting the temperature that provides the optimum results for the detector under consideration. Currently, the operating temperature of infrared detectors ranges from 1.2° K for impurity-activated indium antimonide, to ambient room temperature of 300° K, and above, for certain types of lead salts. The major requirements of a detector cooling system are: long operating time, stable temperatures, light weight, small size, and maximum reliability. Baths of liquefied gases are the most frequently used method of providing low temperatures because of their simplicity of design and compactness [1].

12.1.1. Limitations of Fluid Cooling. The temperature ranges available using baths of liquefied gases are shown in Fig. 12-1. In principle, a liquefied gas can provide constant temperature from the triple point to the critical point; however, except for

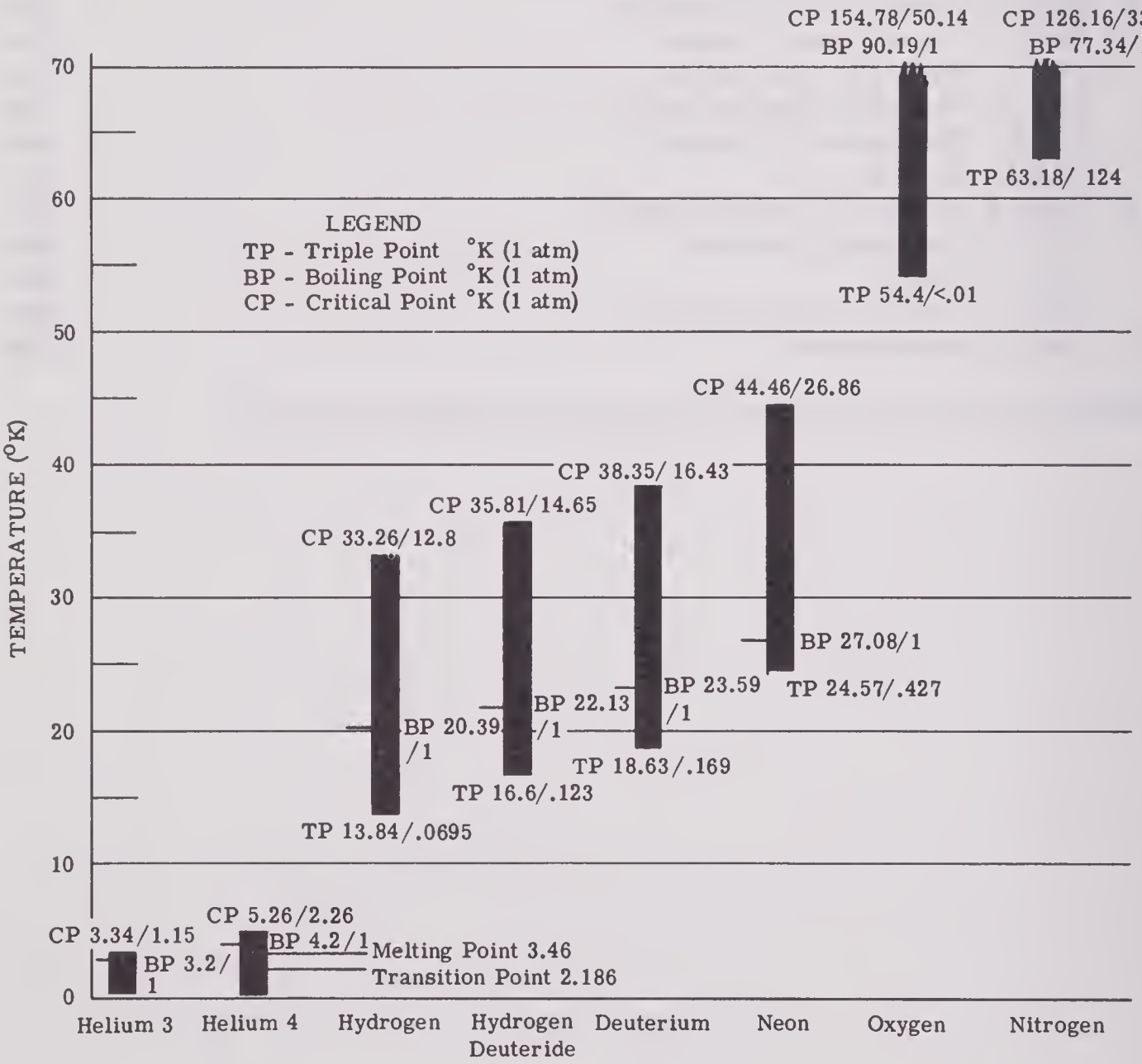


FIG. 12-1. Temperature range of selected low-temperature liquids.

special applications, the boiling point is the usual upper limit. As may be seen from Fig. 12-1, liquefied gases do not exist over the entire temperature range; the lowest temperature at which it is practical to use a liquid bath is 0.3°K, and gaps exist between 4° and 14°K, and between 30° and 55°K. When a gas is above its critical point, it can be made to cool at a temperature below its inversion temperature by adjusting the gas flow to balance the refrigeration produced against the heat load. Temperature stability is sacrificed in operation above the critical point because the heat capacity of gas, as compared to its liquid state, is small. Variation in the heat load therefore creates variation in temperature.

The temperature of the liquid bath can be varied, as shown in Fig. 12-1, by changing the pressure above the liquid by means of a pump and throttle valve. The stabilization of a temperature by holding constant the pressure above the liquid is directly related to the vaporization-temperature history of the liquid. Within the range from the triple point to the critical point, any desired bath temperature can be maintained by holding the pressure constant. At the same time the pumping rate is adjusted to remove precisely the amount of gas vaporized by heat leak into the bath. The cryogenic data for most gases are given in Table 12-1.

Eventually, all baths of liquefied gases boil away because of heat leaks caused by the processes of conduction, convection, and radiation. These processes usually operate simultaneously; however, it is often possible to reduce to a negligible amount the contribution of all but that due to conduction along the solid supports, leads, piping, glass walls, and the insulation itself. The properties of some selected, highly efficient insulating materials suitable for use at low temperatures are given in Table 12-2.

TABLE 12-1. CRYOGENIC DATA [2].

	Boiling Point 1 Atm (°K)	*Melting Point 1 Atm* (°K)	*Liquid Density at bp* (kg/ml $\times 10^{-3}$)	*Gas Density at 273°K & 1 Atm* (kg/ml $\times 10^{-3}$)	*Vapor Density at bp* (kg/ml $\times 10^{-3}$)	*Vapor Pressure Solid at mp* (mm)	*Heat of Vapor at bp* (joules/ kg $\times 10^{-3}$)	*Heat of Fusion at mp* (joules/ kg $\times 10^{-3}$)	*Critical Temperatures* (°K)	*Critical Pressure* (atm)	*Critical Volume* (liter/kg $\times 10^{-3}$)
He^3	3.2	(25 atm)	–	–	–	–	–	–	–	–	–
He^4	4.2	(29 atm)	0.125	0.1785	17.0	–	20.5	4.183	5.2	2.26	0.0144
H_2	20.39	13.98	0.071	0.0899	1.286	54.0	44.8	58.15	33.19	12.98	0.03321
D_2	23.6	18.7	0.173	0.167	2.58	12.8	286	50	38.3	16.2	0.0142
T_2	25.1	21.6	–	–	–	188	–	–	43.7	20.8	0.0089
Ne	27.2	24.47	1.2	0.901	9.5	323	87	16.72	44.5	26.8	0.002
N_2	77.37	63.4	0.808	1.250	4.415	96.5	199	25.52	126.1	33.5	0.00321
CO	81.6	68	0.812	1.186	–	–	213.5	29.27	133.8	35	0.0032
F_2	85.24	53.6	1.513	1.71	–	0.1	171.5	13.4	144.8	55	–
A	87.4	83.6	1.391	1.78	5.03	516	162.7	28.05	150.8	48	0.0019
O_2	90.1	54.9	1.14	1.43	4.75	2	212.5	13.8	154.1	50.1	0.0023
CH_4	111.7	90.7	0.425	0.72	1.76	71	581	60.25	190.5	45.8	0.008
Kr	120.3	116	2.4	3.75	8.33	550	108	16.3	209.3	54.5	–
$R_{14}CF_4$	145.14	89.5	1.62	–	7.2	–	134.8	–	227.5	37	–
O_3	161.3	80.5	1.46	2.14	–	–	316	13.8	261.1	54.6	0.00306
Xe	165.3	150.5	3.1	5.93	9.77	615	96.25	119.1	290	58	0.00086
C_2H_4	169.3	104	0.578	1.19	2.08	–	481	148.5	282.8	50.9	0.0045
N_2O	183.6	183	1.23	1.84	–	658	250.5	95	309.7	71.7	0.0022
C_2H_6	184.8	90	0.562	1.28	0.32	–	490	96.25	305	48.8	0.0048
C_2H_2	189.1	191.2	0.623	1.09	–	–	916	–	309	62	–
$R_{13}CClF_3$	192	91.6	1.505	–	7.9	–	146.4	179.9	302	3.9	–
CO_2	194.6	215.7	1.51	1.87	–	–	574	71.6	304.5	73	0.0022
C_3H_6	226.1	77.5	0.604	1.78	–	–	439.5	–	365	45	–
$R_{22}CHClF_2$	232.5	113	1.414	21.3	4.65	–	235	35.15	369	48.7	–
NH_3	239.8	195	0.683	0.77	0.898	45	1363	–	405	111.2	0.0042
$R_{12}CCl_2F_2$	243.1	118	1.488	17.7	6.25	–	167.2	–	384	39.6	–
CH_3Cl	249.4	–	0.993	5.93	2.56	–	427	–	–	–	–
SO_2	263.1	198	0.80	4.49	3.2	–	388	–	430	77.7	0.002
C_4H_{10}	272.5	–	–	2.53	–	–	–	–	426	36	–
$R_{11}CCl_3F$	296.8	162.7	1.48	2.47	5.93	–	237.3	–	471	43.1	–
C_3H_6	230.8	85.9	0.595	1.92	2.08	–	342	79.9	370	42	–

TABLE 12-2. PHYSICAL PROPERTIES OF SELECTED CRYOGENIC INSULATION [2,3].

Insulation	*Pressure* (mm Hg)	*Temperature Range* (°K)	*Mean Thermal Conductivity* (μw cm^{-1} °K)
Laminae (Cryenco)	10^{-4}	300-77	0.5-2
Opacified Silica Aero Gel	10^{-2}	300-77	2-7
Silica Aero Gel	10^{-3}	76-20	2
Silica Aero Gel	10^{-2}	300-77	20-25
Perlite	10^{-2}	300-77	6.5-11
Perlite – 30 mesh	760	300-77	330
Perlite – 80 mesh	$<10^{-3}$	300-77	10.5
Fiber-Type Glass Fiber	10^{-3}	422	5.8
Heat-Felted Glass Fiber (AA Fiber)	10^{-2}	300-77	7.1
Laminated Type (Linde SI-4)	–	300-88.5	0.4
NRC-I	10^{-5}	300-20	0.9-1
0.008-in. glass paper, 0.0023-in. aluminum foil – 55 shields/in.	$<10^{-5}$	300-20	0.4

12.1.2. Cooling System Design Criteria. The main variables entering into the design of the particular infrared cooling system are:

1. Refrigeration load
   - *a.* Radiation load on cooled surfaces
   - *b.* Conduction leaks through mechanical supports
   - *c.* Conduction leaks through electrical leads
   - *d.* Cell bias power
2. Operating time at rated load
3. Standby time with no refrigerant flow
4. Cooling temperature and tolerances
5. Environmental operating conditions
6. Weight and space requirements
7. Detector cell configuration and dimensions
8. Length and type of feed lines (especially critical in liquid transfer coolers)

Once this information has been obtained, a choice can be made as to the type and the capacity of the cooling system.

12.2. Types of Detector Cooling Systems

The five types of detector cooling systems are: direct contact, Joule-Thomson, expansion-engine, thermoelectric, and magneto-thermodynamic. Of the five, three are primarily mechanical, one is electrical, and one is magnetic. The three mechanical cooling systems can be further classified as open or closed cycle, depending on whether the evaporated coolant is vented to the surroundings or recycled. Magneto-thermodynamic cooling, although a promising process for the generation of temperatures approaching a few microdegrees Kelvin, is not currently used for detector cooling because the state of the art is such that constant low temperatures cannot be maintained for any length of time. Table 12-3 lists the characteristics of the various types of

TABLE 12-3. INFRARED DETECTOR COOLING SYSTEMS CHARACTERISTICS

Cooler	*Type*	*Temperature Range* (°K)	*Cooling Capacity* (watts)	*Power Input* (watts)	*Cooldown Time* (min)
Direct Contact	Integral	4.2-77	0.05-10	None[1]	0-15
	Liquid feed	4.2-77	0.05-10	None[1]	0.05-15
Joule-Thomson	Single stage, open cycle	20-80	0.01-5	None[1]	1-5
	Multiple stage, open cycle	4.2-27	5-20	None[1]	5-20
	Single or multistage closed cycle	4.2-77	1.0-30	250-1000	5-30
Expansion Engine	Piston-Regenerator	30-300	0.01-1.0	400-800	3-10
	Displacer-Regenerator	30-300	1-10	200-300	5-10
	Turbine	30-300	1.5-5	400-600	5-10
Thermoelectric	Single stage	250-300[3]	0.01-0.2	0.5-3	2
	Cascaded multiple stages	195-250[3]	0.01-0.02	1-5	2

TABLE 12-3. INFRARED DETECTOR COOLING SYSTEMS CHARACTERISTICS (*Continued*)

Cooler	*Design Features*	*Design Limitations*
Direct Contact	Excellent reliability, light weight, small bulk, low pressures[2]	Short operating time limited by liquid capacity; standby time limited by evaporation rate
	Excellent reliability, remote cooling, light weight, small bulk, simplified installation problems, low pressures[2]	Short operating time limited by liquid capacity; standby time limited by evaporation rate, losses in transfer lines
Joule-Thomson	Remote cooling, light and small cooling head, simplified installation problems[2]	Easily clogged; requires high pressures and flow rates; high leakage and contamination; short operating time limited by tank capacity; only portion of tank used because of pressure drop
	Low temperatures obtained, remote cooling, light and small cooling head, simplified installation problems[3]	Needs precooling with separate gas cycle; easily clogged; requires high pressures and flow rates; high leakage and contamination; short operating time limited by tank capacity; only portion of tank used because of pressure drop
	Continuous duty, long operating time, low temperatures obtained, remote cooling, light and small cooling head, simplified installation problems	Requires highly loaded, high-pressure, noncontaminating compressor; poor life and nonreliable; needs precooling for low temperatures; easily clogged; high leakage and contamination
Expansion Engine	Continuous duty, long operating time, low pressures	Requires noncontaminating compressor; great wear and constant friction in expander; high leakage around expander valves; low regenerator efficiency, microphonics
	No valves required, continuous duty, long operating time, low pressures, low friction with gas bearings	Great wear and constant friction between piston, displacer, and cylinder; poor compression; low regenerator efficiency; microphonics
	No valves required, continuous duty, long operating time, low pressures, low friction with gas bearings	Requires high flow rate, means of removing work done on turbine, separate compressor; strong mechanical construction required because of high speed
Thermoelectric	Excellent reliability, static operation, light weight, very small, long life	Relatively high-temperature operation; low cooling capacity; hard to obtain operating voltages and currents
	Excellent reliability, static operation, light weight, very small, long life, lower temperatures	Relatively high-temperature operation, very low cooling capacity

[1]May require 20 to 100 w for storage-tank heater and for control system.
[2]Size and weight of complete system depends on storage-tank capacity and operating time.
[3]Lower temperatures will be available when more efficient thermoelectric materials under investigation are developed.

detector cooling systems presently available commercially. Manufacturers of detector coolers are listed at the end of this chapter. The temperature ranges of these systems are shown in Fig. 12-2. For the mechanical cooling systems, the temperature range is dependent on the fluid used. The range of the common liquids is shown; however, other fluids and custom-built designs can broaden these limits.

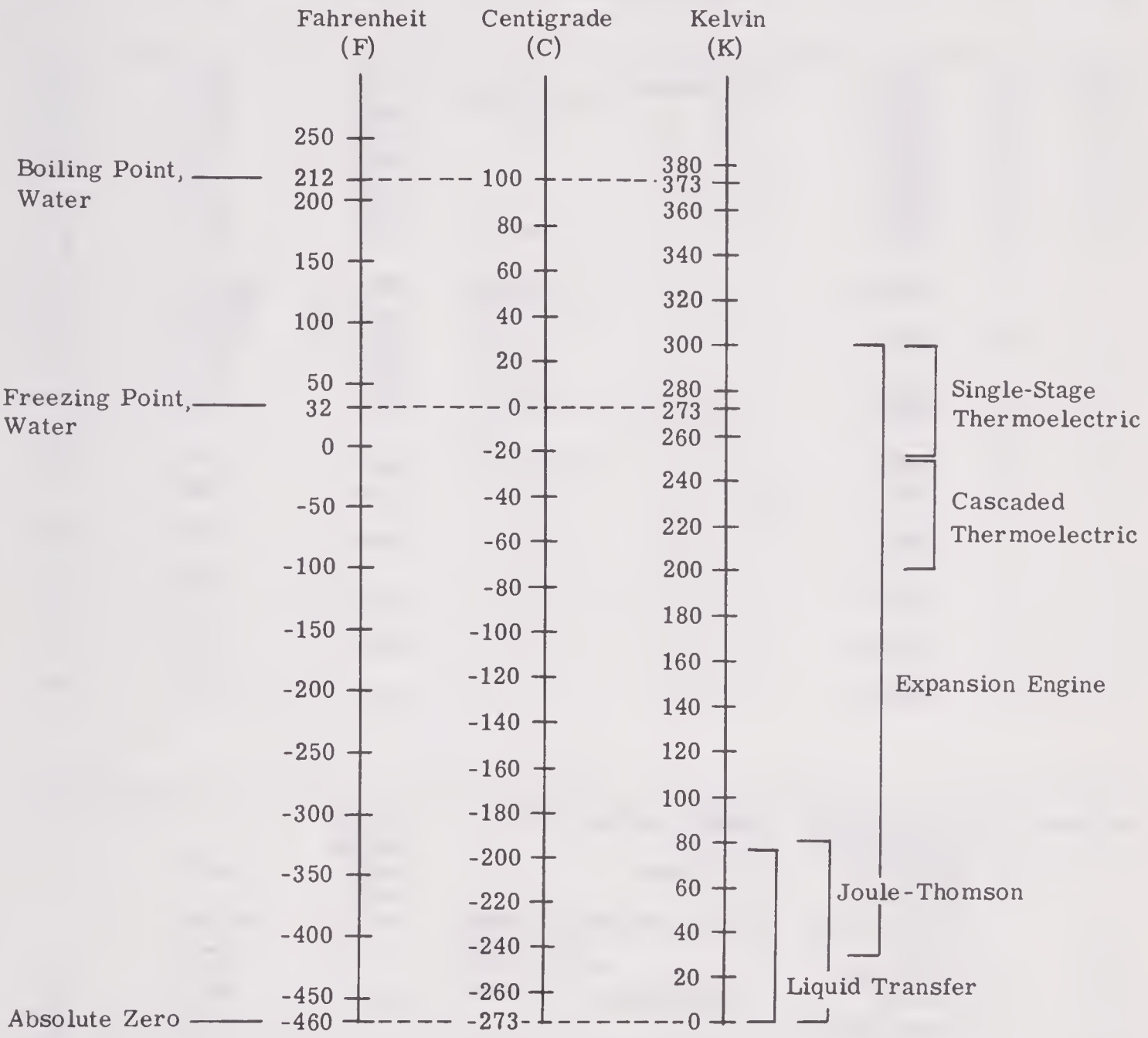


FIG. 12-2. Temperature range of commercially available infrared detector cooling systems.

12.2.1. Direct-Contact Coolers. There are two types of direct-contact coolers: integral and liquid feed. The integral cooler is the most common of all cooling systems. In its simplest form, it consists of nothing more than a detector dewar into which the coolant is poured. The liquid-feed cooler is a more involved version of the basic direct-contact cooler. The coolant liquid is fed from a liquid storage tank through transfer lines to the cooling head. The principal advantage of the liquid-feed cooler is its installation flexibility; the cooling head can be mounted remote from the liquid-storage container. Table 12-4 lists specific models of commercially available direct-contact coolers.

TABLE 12-4. LIQUID-TRANSFER COOLERS.

Manufacturer	*Model*	*Type*	*Coolant*	*Capacity* (liters)	*Operating Temperature* (°K)	*Cooldown Time* (min)
AiResearch	134338-1	Liquid feed	N_2	5.0	77	*a*
	134642	↓	N_2	1.5	77	3
	134548			1.0		3
ITT	–	Liquid feed	N_2	1.83	77	<2
Linde	LNI-1	Integral	N_2	0.174	77	*a*
	LNI-13			0.282		
	LNI-18			0.198		
	LNI-3			0.902		
	LNI-4			0.174		
	LNI-5			0.209		
	LNI-12		Ne	0.777	27	
	LNI-15		Ne/He	0.395	Ne=27	
			N_2	0.062	He=4	
	LNI-28		Ne/He	1.33		
			N_2	0.450		
	LNI-9		N_2	1.181	77	
	LNF-2	Liquid feed	N_2	1.524		*b*
	LNF-3			0.456		
	LNF-4			3.159	77-166	
	LNF-5			3.159	77	
	LNF-6			0.519		
	LNF-12			0.405		
	LNF-13	↓	↓	1.128	↓	↓
Raytheon	QKN 748 & 1003	Integral/liquid feed	N_2	0.00005	77	*a*
	QKN 884 & 1004	↓		0.00119	↓	
	QKN 1204	Integral		0.045	65-77	
	QKN 1205		↓	0.007	↓	
	QKN 1206, 1207 & 1208	Integral/liquid feed	Ne/He N_2	0.05/0.5	Ne=27 He=4	
	Storage dewar	↓	↓		↓	
	Transfer	Liquid feed	N_2	0.147	77	2
	Pumped transfer	↓	↓		68	2
SBRC	LNI-E	Liquid feed	N_2	1.5 also 1 & 3	77	<1
Hughes Aircraft Co.	DP-099	Liquid feed	N_2	1.356	77	<3
	DP-001	Liquid feed	N_2	1.163	77	<3
	AP-111	Liquid feed	N_2	3.172	77	<3

TABLE 12-4. LIQUID-TRANSFER COOLERS (*Continued*).

Standby Time (hr)	*Evaporation Rate* (kg/hr)	*Operating Time* (hr)	*Method of Filling*	*Length* (m)	*Diameter* (m)	*Weight*	
						Empty (kg)	*Full* (kg)
24	0.0081	30-50	Pressure lines	0.254	0.28	3.4	7.5
24	0.0136	6	↓	0.406	0.127	1.59	2.94
24	0.00906	3.5		0.178	0.1525	2.27	3.08
83	0.0181	0.3	Pour or pressure	0.292	0.1568	2.72	4.2
8.5	0.0168	8.5	Pour	0.1015	0.076	0.202	0.344
8.0	0.0281	8.0	Insulated fill lines	0.146	0.1175	1.81	2.02
8.0	0.0199	8.0	Pour	0.133	0.0761	0.226	0.386
29.0	0.025	29.0	↓	0.33	0.089	0.78	1.455
10.0	0.014	10.0		0.1365	0.0794	0.163	0.304
9.0	0.0186	9.0		0.2255	0.0761	0.286	0.454
8.0	0.0785	8.0	Insulated fill lines	0.203	0.1175	2.04	2.66
Ne=100[c]	Ne=0.00454	Ne=100[c]	Insulated fill lines,	0.277	0.089	0.955	*d*
He=3[c]	He=0.0168	He=3[c]	pour for shield				↓
Ne=190[e]	Ne=0.0081	Ne=190[e]	↓	0.394	0.127	1.585	
He=5.75[e]	He=0.029	He=5.75[e]					
2.0	0.476	2.0	Pour	0.254	0.1142	1.61	2.6
24.0[f]	0.0172	*h*	Pressure	0.464	0.1142	1.36	2.58
24.0[g]	0.0059	↓	Pour or pressure	0.254	0.089	0.905	1.275
204	0.0122		Pour	0.222×0.1715 ×0.343	–	3.17	5.61
204	0.0122		↓	0.343	0.1525	1.27	3.81
<1	0.666			0.305	0.089	0.294	0.712
<1	0.870	↓		0.306	0.089	0.408	0.735
24	0.02675	3	↓	0.254	0.1015	1.585	2.5
			Pour	0.0444	0.0127	0.011	
			↓	0.0761	0.0127	0.020	
		0.75		0.0761	0.019	0.0113	
		1.5		0.0761	0.019	0.0113	
		8	Insulated fill lines, pour for shield	0.4165	0.1142	3.4	3.74
		8	Pour or pressure	0.353	0.1142	2.78	3.12
48[i]	0.2175[i]	5[i]	Pressure lines	0.508	0.1525	2.26	3.44
48[i]	0.277[i]	4[i]	↓	0.711	0.203	5.9	7.08
60		8[j]	Pressure lines	0.4835	0.127	2.04	
24	0.014	4[j]	Insulated fill lines	0.28	0.135	2.04	3.18
24	0.018	4[j]	Insulated fill lines	0.31 height 0.29 width 0.09 depth		3.22	4.22
24	0.028	8	Insulated fill lines	0.42	0.153	3.98	6.58

[a] Cooled concurrently with filling.
[b] Depends on size and construction of cooling head.
[c] Based on a 0.09-w detector and electrical lead heat load.
[d] Sum of empty weight, weight of shield liquid, and weight of inner cell liquid.
[e] Based on 0.02-w detector and electrical lead heat load.
[f] Evaporation rate of 0.408 kg/day, 0.816 kg of liquid remaining after 24-hr standby.
[g] Evaporation rate of 0.1405 kg/day, 0.266 kg of liquid remaining after 24-hr standby.
[h] Dependent on flow rate and heat load.
[i] Based on the use of a QKN 1004 dewar.
[j] Based on a 1/2-w heat load.

TABLE 12-4. LIQUID-TRANSFER COOLERS (*Continued*).

Manufacturer	*Detector Location*	*Operating Attitude*	*Remarks*
AiResearch	External	Vertical	Flow controlled by heat leak pressure buildup.
	↓	Horizontal	Flow controlled by heat leak pressure buildup.
		Vertical	Flow controlled by heat leak pressure buildup.
ITT	External	Vertical	Flow controlled by heat leak pressure buildup; transfer-head cooling tube 0.419 m long, 0.00635 m i.d., weight 0.0396 kg.
Linde	Side	Vertical	Separate fill and vent ports.
	Bottom	Vertical	Separate fill and vent valves.
	End	Horizontal	Combination fill and vent port.
	Tail, side	Horizontal	Separate fill and vent ports; 0.1208-m tail.
	Tail, side	Horizontal	Combination fill and vent port; 0.0635-m tail.
	Tail, bottom	Vertical	Separate fill and vent ports; 0.092-m tail.
	Bottom	Vertical	Temperature is pressure dependent; cools mosaic detectors.
	Bottom	Vertical	Double cell for Ne or He with N_2 shield.
	Bottom	Vertical	Double cell for Ne or He with N_2 shield.
	End	Horizontal	Single cell; used to cool vidicon tubes.
	External	Horizontal	Flow controlled by electric heater pressure buildup; requires 24-v dc, 115 w.
	↓	Vertical	Flow controlled by orifice valve.
		Vertical	Flow controlled by bridge circuit and valve regulating pressure buildup.
		Vertical	Flow controlled by valve on dewar vent or temperature control panel.
		<60° from vertical	Flow controlled by evaporation rate of liquid in cell dewar.
		<60° from vertical	Flow controlled by evaporation rate of liquid in cell dewar.
		Vertical	Orifice-control or temperature-control devices to regulate flow rate, absolute or gauge pressure-relief valves, with optional pressure buildup heaters for system pressure regulation.
Raytheon	Bottom	Horizontal to vertical	Small metal dewar with integral detector.
	↓	↓	↓
		<55° from vertical	Double cell for Ne or He with N_2 shield.
		<55° from vertical	Double cell for Ne or He with N_2 shield.
		Horizontal	Flow controlled by orifice valve.
		Horizontal	Flow controlled by orifice valve and pump rate.
SBRC	External	Horizontal	Flow controlled by orifice valve; 0.00519-m or 0.00828-m diameter cooling head; weight 0.0142 kg.
Hughes Aircraft Co.	External	Horizontal	Flow controlled by temperature-control device, relief valve, and heater.
	External	Horizontal	Flow controlled by orifice, relief valve, and heater.
	External	Horizontal	Flow controlled by absolute pressure relief valve and temperature-control device; system designed for high ambient temperature environment.

12.2.1.1. *Integral Coolers.* The integral cooler consists of a detector in direct thermal contact with a supply of liquid coolant (Fig. 12-3). The detector is integrally mounted in a dewar that serves both as the detector mount and the liquid container. When a solid coolant such as dry ice is used, sticks of the coolant are inserted into the coolant well. Thermal contact between the solid CO_2 and the walls of the coolant well is ensured by mixing the solid with a low-freezing-point liquid such as acetone. A basic limitation of the direct-contact cooler is its operating attitude. In order to keep the coolant in direct contact with the detector, the dewar must be maintained in an essentially vertical position. For airborne and tracking instrument applications where the detector is moved through 360° of an arc in both horizontal and vertical planes, thermal contact between the coolant and the detector is maintained regardless of the dewar attitude. This is accomplished by using copper conducting plates that are spring loaded so that they remain in contact with the coolant.

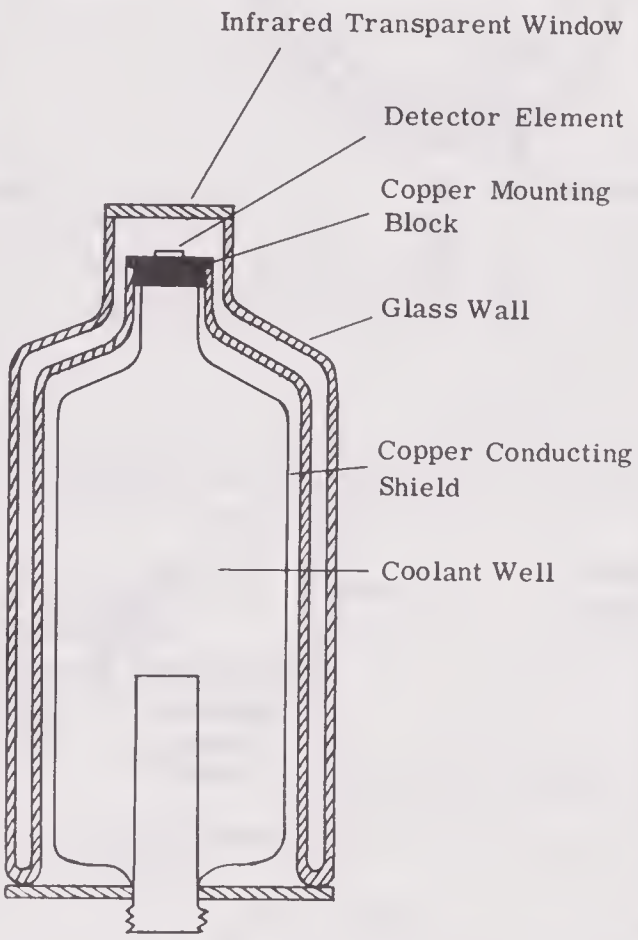


FIG. 12-3. Direct-contact cooler.

12.2.1.2. *Liquid-Feed Coolers.* The liquid-feed cooler consists of an insulated liquid-storage container, transfer lines, a cooling head, and the necessary controls (Fig. 12-4). The transfer mechanism is either gravity or gas pressure. The gas pressure to force the liquid from the storage container to the cooling head originates from the natural pressure build-up due to thermal leakage into the storage container, or from the residual pressure of the filling operation. In cases where the natural pressure build-up is not sufficient, or better regulation is required, a small pressure-regulated electrical heater is placed in the storage container to evaporate the required amount of liquid (see Fig. 12-5). The flow of fluid is usually self-limiting to provide operation over a wide range of differential pressures, using only on-off control, and to prevent flooding of the cooling head. As the detector cell cools, the liquid in the cell evaporates and vents through the adjustable orifice-flow control valve, which sets the pressure differential between the tank and the cooling head to a value that maintains a constant rate of flow

of liquid into the cell. The pressure-control relief valve regulates the pressure exerted on the liquid, and also acts as a relief valve to vent the storage container in case of malfunction. With the addition of a small gas liquefier to the low-pressure cell vent line, the liquid-feed cooler becomes a closed-cycle system capable of long standby and operating times.

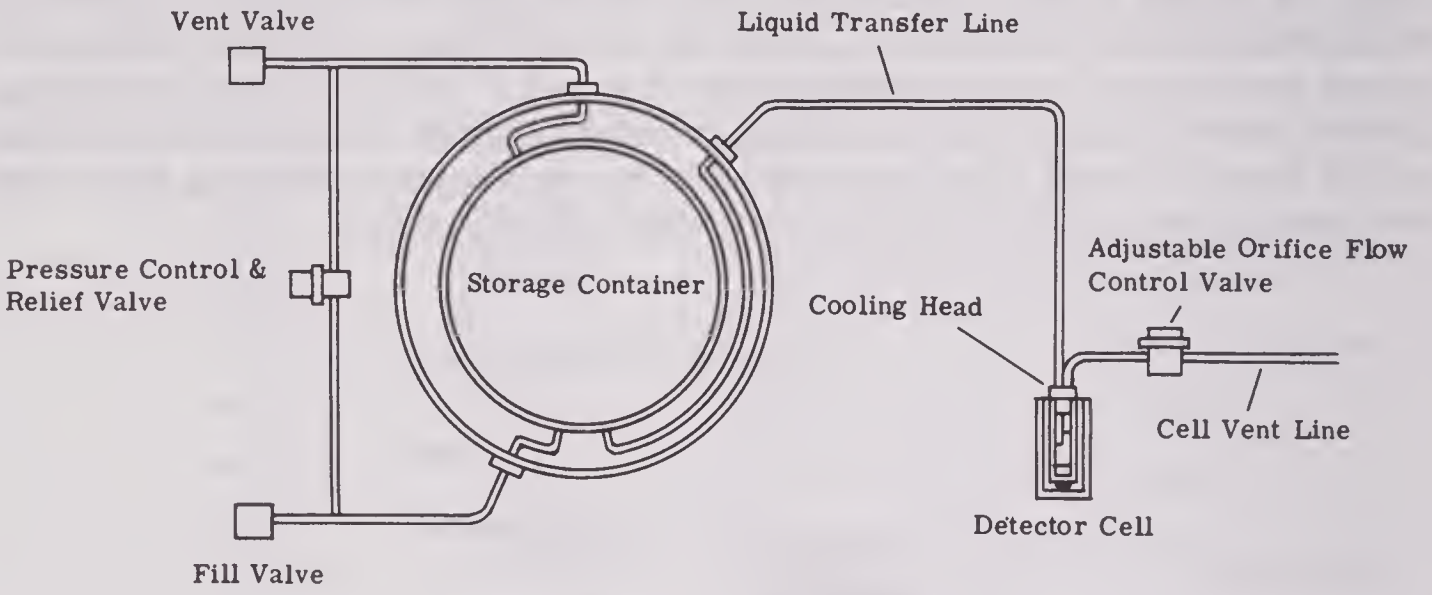
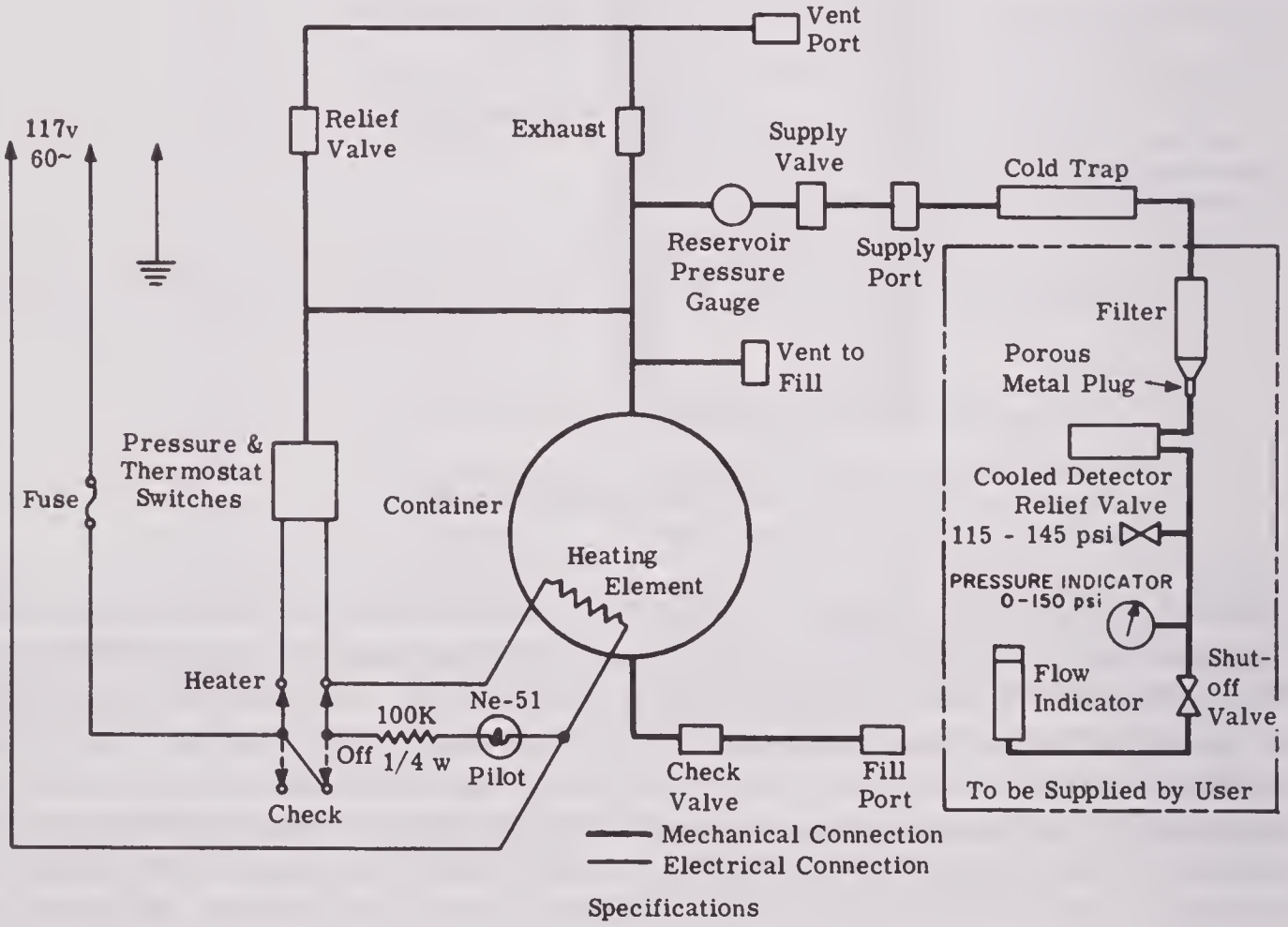


FIG. 12-4. Liquid-feed cooler.



Specifications

Liquid Nitrogen Charge	5.0 liters	Operating Time	5.5 hr at 5 lpm after 2 hr standby, 2.25 hr at 12 lpm at 2000 psi
Output Pressure	2000 psi		
Heater Off Pressure	2300 psi min		
Pressure Buildup Time	10 min to 1500 psi	Weight	22.125 lb uncharged
		Dimensions	14.812 × 12 × 13.5 in.

FIG. 12-5. Barnes nitrogen pressure generator diagram [4].

The usual mechanism of liquid transfer is a two-phase flow known as Leidenfrost transfer. When small quantities of a low-temperature liquid are passed through a warm tube, some of the liquid evaporates to form a gaseous skin that keeps droplets of the liquid insulated from the walls of the tube sufficiently well so that small quantities of the liquid can be efficiently transferred. The loss of liquid in the transfer process increases with transfer line length, typically 75% for a line 1 ft long. Figure 12-6 shows a plot of transfer efficiency *versus* cooling load for several lengths of uninsulated transfer lines. Graphs of total system weight *versus* total operating time and *versus* required liquid capacity for a typical liquid-nitrogen detector cooling system are shown in Fig. 12-7. Using these graphs, a single-search detector load of 1 w operating for 100 hr would require a system weighing 7.6 lb with a liquid-storage container of 5.9-liter capacity.

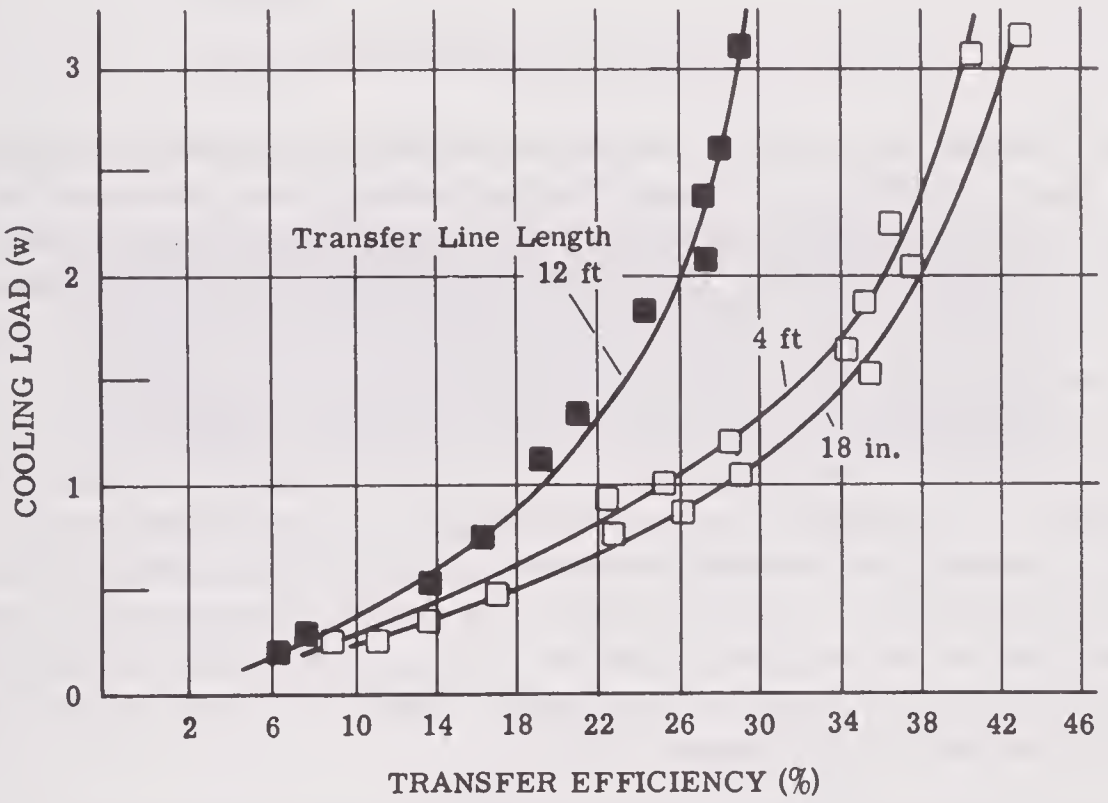


FIG. 12-6. Transfer efficiency for liquid N_2 through uninsulated tubes [1].

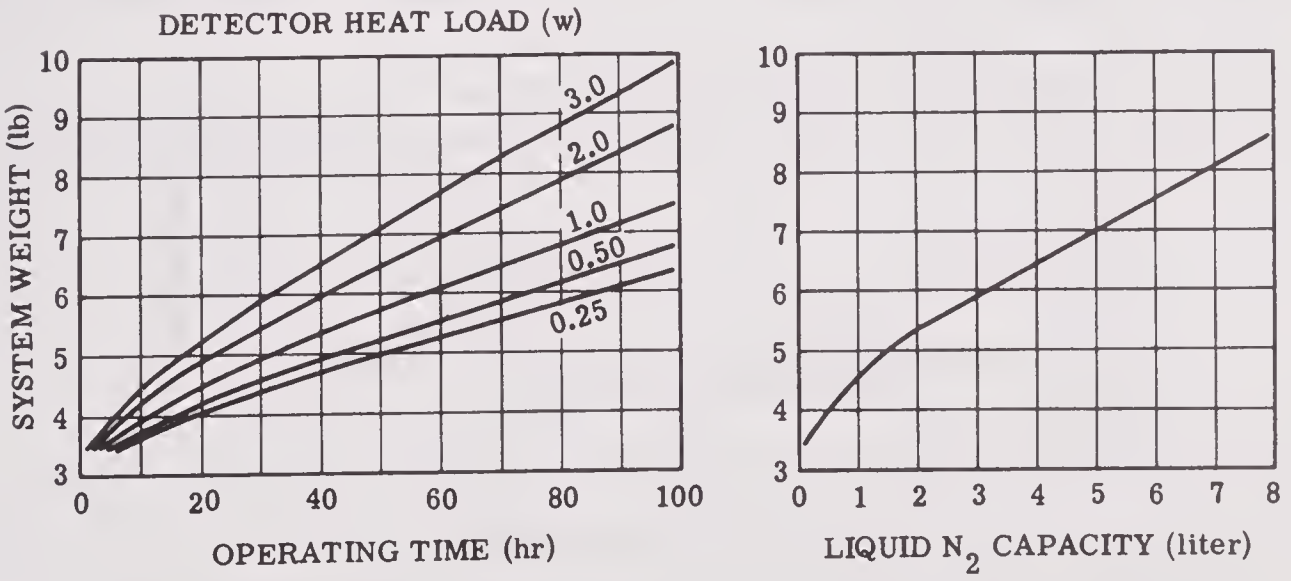


FIG. 12-7. System dry weight *versus* operating time and storage tank capacity [5].

12.2.2. Joule-Thomson Coolers. The Joule-Thomson or cryostat cooler is based on the Joule-Thomson effect of cooling caused by the adiabatic expansion through an orifice of a gas that is initially below its saturation temperature. The expanded gas, thus cooled, is passed back over the incoming gas to cool that, which results in regenerative cooling. The process continues until liquid begins to form at the orifice to produce a bath of liquid at the boiling temperature of the gas. The Joule-Thomson cooler consists of a finned tube in the form of a coil, an orifice, and orifice cap, and an outer shield or coil (Fig. 12-8). The finned tube is made of very small inside-diameter

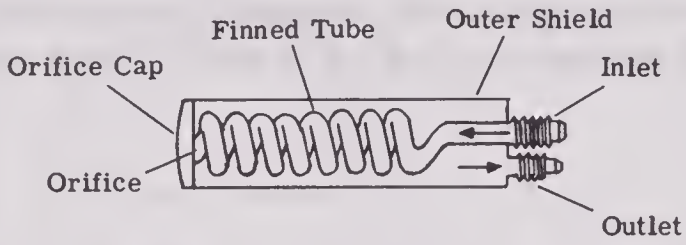


FIG. 12-8. Construction of Joule-Thomson cooler.

tubing to provide the large ratio of surface area to volume necessary for effective heat exchange. Dust particles and all traces of higher-freezing-point gases must be excluded from the gas entering the cooler; otherwise the finned tube will clog or freeze up after a short period of operation. The orifice is usually surrounded by an orifice cap that returns the expanded gases back along the finned tube. This cap must be maintained in close thermal contact with the dewar or detector to ensure efficient heat transfer. The outer shield is formed in a number of ways. A separate return tube can be coiled around the finned tube, the finned tube can be placed in another length of tubing of greater diameter, or the fined tube can be placed in a tight-fitting container so that the expanded gas must flow through the channels formed by the space between adjacent turns of the finned tubing. With the addition of a small noncontaminating compressor, the expanded gas can be recycled to provide continuous operation for long periods of time without the need for gas replenishment. Table 12-5 lists specific models of commercially available Joule-Thomson coolers.

12.2.2.1. *Gas-Supply Coolers.* In an open-cycle Joule-Thomson cooler, bottled gas is the usual source of supply, although high-pressure gas containers are frequently used. A high-pressure regulator is required to maintain the gas pressure at approximately 1500 psi since both the flow rate and cooldown time are pressure dependent (Fig. 12-9). Most pressure regulators that cover the range of 1000 to 3000 psi, and

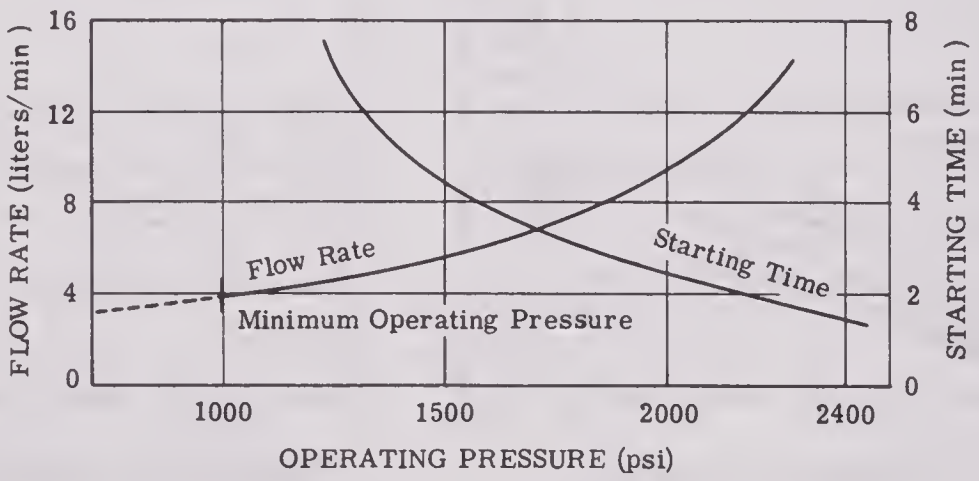


FIG. 12-9. Flow rate required to liquefy nitrogen and cooldown time as a function of gas pressure for SBRC P/N 9114 cryostat in style No. 120-1 detector. Total detector flask heat load of 0.23 w and ambient temperature of 295° K [6].

accommodate normal flow rates up to 10 liters per minute (standard temperature pressure), can be used. The pressure regulator must be capable of reducing the pressure to counteract the tendency of the flow to increase, sometimes reaching a value twice as much as the initial flow, as the temperature drops. Commercially available high-purity gas with a low water content is required for operation in the cooler. To ensure that all undesirable particles are removed, and that the water content is sufficiently low, a dryer and a filter can be inserted in the high-pressure line ahead of the finned tube. A chemical dryer containing an absorbent material removes the excess moisture; it is followed by a porous or a sintered metal filter to remove any floating particles. To obtain even greater gas purity, a cold trap consisting of a coiled tube immersed in a bath of dry ice and acetone or liquid nitrogen is inserted after the dryer and filter (Fig. 12-10). In operation, water vapor and gases, having a freezing point above that of the coolant bath, will condense on the walls of the tubing at a rate dependent on the initial purity of the gas being dried.

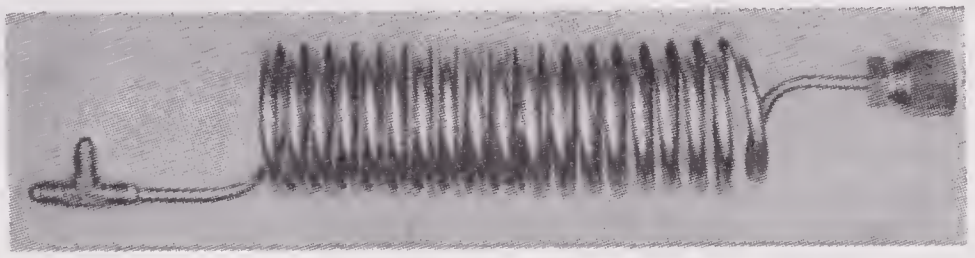


FIG. 12-10. High-pressure cold trap with metal filter [7].
(*Copyright ITT, Farnsworth Division*).

12.2.2.2. *Liquid-Supply Coolers.* The simplest way to obtain high-purity dry gas is to feed the Joule-Thomson cooler from a storage container of liquid coolant. A major problem of the liquid supply is the evaporation of liquid during standby; however, this supply is still lighter and smaller than a compressed gas supply of the same capacity. A diagram of a typical liquid supply or nitrogen pressure generator is shown in Fig. 12-5. The pressure-regulated heater inside the storage tank evaporates the liquid to provide the proper gas pressure. A cold trap, a filter, and a porous metal plug remove undesirable particles and all traces of higher-freezing-point gases.

TABLE 12-5. JOULE-THOMSON COOLERS

Manufacturer	*Model*	*Cycle*	*Coolant*	*Operating Pressure* (atm)	*Flow Rate* (ℓ/min)	*Op. Temp.* (°K)	*Cooling Capacity* (w)
AiResearch	Joule-Thomson	Closed	N_2	102.1		77	1
Air Products Inc.	AC-2-109A	Open	H_2[a]	122.4	10.5	20-75	0.75
			N_2	152.9	9.9		
	AC-2-109B	Open	H_2	105.5	9.1	20-75	1.5
			N_2	122.4	27.2		
	AC-2-109C	Open	H_2	102.1	29.7	20-75	4.0
			N_2	142.7	31.2		
	AC-2-109D	Open	H_2	102.1	21.0	20-75	6.0
			N_2	136.0	36.2		
	NOTS-Side-winder 1C	Open[b]	N_2	81.7-170.3	5.5	80	5-8
	IOR-OON	Closed	N_2	68.0	8.5	80	1
	2C	Closed	N_2	54.4-81.7	8.5-22.6	80-100	1-5
	Two-fluid cascade	Closed	H_2 N_2	34.0-81.7	17-31	30	0.5-1
ITT	600[c]	Open	N_2	115.6 149.5 max	2-6 min	77	0.01
	600[c]	Open	N_2	115.6 149.5 max		77	0.01

[a]Available for operation with other gases and different pressure and flow rates.

[b]Similar Joule-Thomson coolers for use with several of the gases, argon, neon, hydrogen, and helium are available.

[c]Compressed N_2 supply consisting of 8.2-liter tank, solenoid-operated valve, filler, pressure gauge, and quick disconnect available. Operates for 1-1/2 hours, gas bottle charging cart available. Cryostats recommended for use with an ITT detector cell dewar.

TABLE 12-5. JOULE-THOMSON COOLERS (*Continued*)

Cooldown Time (min)	*Construction*	*Heat Exchanger*			*Compressor*	*Required Accessories*
		Length (m)	*Diameter* (m)	*Weight* (kg)		
	Bundle-type counterflow	0.203	0.089	0.454	Two stage, nonlubricated piston-type, ac motor driven, 115 v, 130 w, 60 cps, 1 ϕ, 0.432 × 0.178 m diameter, 9.07 kg.	
5.5	Two-stage cascade,	0.114	0.00635	0.454	Not Used	*g*
19	concurrent flow	0.159	0.00955	0.454	Not Used	*g*
12	Two-stage cascade,	0.21	0.0127	0.454	Not Used	*g*
22	concurrent flow	0.33	0.019	0.454	Not Used	*g*
0.67-1	One stage concurrent flow	0.0457	0.00508	0.454	Not Used	*g*
30	One stage, finned coiled tubing	0.0762	0.00635		Two stage, nonlubricated piston-type, ac motor driven, 110 v, 250 w, 60 cps, 1 ϕ; or 208 v, 600 w, 400 cps, 3 ϕ, 0.127 × 0.205 × 0.305 m, 7.25 kg.[c]	*g*
10-25	One stage, finned coiled tubing	0.0762-0.152	0.00635-0.0127		Two stage, nonlubricated piston-type, ac motor driven, 110 v, 300-1000 w, 60 cps, 1 ϕ; or 208 v, 300-1000 w, 400 cps, 3 ϕ, 0.127 × 0.205 × 0.28-0.381 m,[d] 7.25-12.7 kg.[c]	*g*
20-30	Two stage, finned coiled tubing	0.152-0.239	0.00955-0.0127		Two nonlubricated, two stage piston-type, ac motor driven, 110 v, 800-2000 w, 400 cps, 3 ϕ; or 208 v, 800-2000 w, 400 cps, 3 ϕ, 0.127 × 0.205 × 0.28-0.381 m,[c] 15.9-22.8 kg.[c]	*g*
20	One stage, coiled tubing	0.114	0.00955		Not Used	*g*
20	One stage, coiled tubing	0.143	0.00476		Not Used	*g*

[c]A 2-stage, 6-cylinder compressor has been developed for use where lower temperatures and greater cooling capacity are required, but it is not recommended for infrared use.

[d]The 2-fluid cascade cooler requires two model 2C units, each with the listed dimensions, which can be packaged together or separately as desired.

[g]Filter absorber system is a required accessory. Contact firm for details.

TABLE 12-5. JOULE-THOMSON COOLERS (*Continued*)

Manufacturer	*Model*	*Cycle*	*Coolant*	*Operating Pressure* (atm)	*Flow Rate* (ℓ/min)	*Op. Temp.* (°K)	*Cooling Capacity* (w)
Honeywell		Open or closed	N_2	136.0 102.1	3 start 4.5 steady	77	1
Hughes Aircraft Co.	JTC-001	Closed	N_2	68.0- 103.6	9 each for 6 J-T heat exchr	77	0.4 for each J-T
	JTC-002	Closed	A	68.0- 103.6	7.5 each for 6 J-T heat exchr	87	0.4 for each J-T
	JTO-001	Open	A	406.5- 116.1	40-10	87	40-10
	JTO-002	Open	A	358.1- 62.9	15-8	87	10-1
Philco	C100	Open	N_2	68.0- 204.2	3 start 9 steady	77	
	C150		N_2			77	
Raytheon	QKN 748/1003	Open	N_2			77	
	QKN 884/1004	Open	N_2	95.2	31		
	QKN 961/1005/ 1007	Open	H_2/Ne N_2			20/28 65-78	
	QKN 1204	Open	N_2			65-78	
	QKN 1205	Open	N_2			65-78	

TABLE 12-5. JOULE-THOMSON COOLERS (*Continued*)

Cooldown Time (min)	*Construction*	*Heat Exchanger*			*Compressor*	*Required Accessories*
		Length (m)	*Diameter* (m)	*Weight* (kg)		
2	One stage, coiled tubing	0.122	0.0042 min		Not Used	*g*
4.5	One stage, finned coiled tubing	0.051	0.0083	0.030	Four stage, nonlubricated, piston type with unique seal separating lubricated and nonlubricated areas. ac motor driven, 208 v, 400 cps, 3 ϕ, 1300 w input; 2 compressors with common drive can be driven hydraulically or electrically; both direct and belt drive; 0.41 × 0.46 × 0.28 m, 30 kg.	
3.0	One stage, finned coiled tubing					
0.15	One stage, integrally finned tubular exchanger	0.032	0.005	0.010	Not Used	*h*
0.5	One stage, integrally finned tubular exchanger	0.051	0.0063	0.015	Not Used	*h,i*
1-1.5	One stage, coiled tubing	0.07	0.0097	0.085	Not Used	*g*
0.75-0.92	One stage, coiled tubing	0.07	0.00524	0.085	Not Used	*g*
	One stage, finned coiled tubing	0.0317	0.00515		Not Used	*g*
		0.0635	0.00515		Not Used	*g*
	Two stage, coiled tubing	0.2135	0.00896		Not Used	*g*
	One stage, finned coiled tubing	0.0635	0.00896		Not Used	*g*
		0.0635	0.00514 min 0.0127 min		Not Used	*g*

[g] Filter absorber system is a required accessory. Contact firm for details.
[h] Storage tank required.
[i] Manifold required.

TABLE 12-5. JOULE-THOMSON COOLERS (*Continued*)

Manufacturer	*Model*	*Cycle*	*Coolant*	*Operating Pressure* (atm)	*Flow Rate* (ℓ/min)	*Op. Temp.* (°K)	*Cooling Capacity* (w)
SBRC	9114	Open	Freon	17.0		192	0.2-1
			A	68.0		80	
			N_2	81.7	4.5	77	
	9174	Open	Freon	17.0		192	0.2-1
			A	68.0		80	
			N_2	81.7	4.5	77	
	9186	Open	Freon	17.0		192	0.2-1
			A	68.0		80	
			N	81.7	4.5	77	
	9185	Closed	Freon	17.0		192	0.2-1
			A	68.0		80	
			N_2	81.7	4.5	77	
	80°K – 0.5 w	Closed	N_2			82	2
	80 K 5 w	Closed	N_2	136.1		80	5
	80 K 2 w	Closed	N_2	102.1		80	2
	82 K .5 w	Closed	N_2	102.1		84	1
Stratos	S-2108	Closed	N_2	199.8	9.06	26	1.0
			H_2	136.0	7.08		
Westinghouse		Closed	N_2	102.1	7.08	90	0.5
		Closed[f]	Ne/N_2			43	

[f] 2-liquid closed-loop cooler and dual-diaphragm compressor under development.

TABLE 12-5. JOULE-THOMSON COOLERS (*Continued*)

Cooldown Time (min)	*Construction*	*Heat Exchanger*			*Compressor*	*Required Acessories*
		Length (m)	*Diameter* (m)	*Weight* (kg)		
<3	One stage, finned coiled tubing	0.042	0.00519	0.00227	Not Used	*g*
<3	One stage, finned coiled tubing	0.0418	0.0083	0.00255	Not Used	*g*
<4	One stage, finned coiled tubing	0.0508	0.00519	0.00227	Not Used	*g*
<4	One stage, finned coiled tubing	0.0478	0.0083	0.00255	Not Used	*g*
5	One stage, finned coiled tubing	0.0508	0.00519	0.02834	4 cylinder piston, ac motor driven, 208 v, 225 w, 400 cps, 3 ϕ, 0.241 × 0.152 m diameter, 4.54 kg. 3-stage piston, ac motor driven, 208 v, 525 w, 400 cps, 3 ϕ, 3.175 × 1.78 m diameter, 7.3 kg. Air cooled with own fan. 3-stage piston, ac motor driven, 208 v, 325 w, 400 cps, 3 ϕ, 3.175 × 1.78 m diameter, 7.3 kg. Air cooled with own fan. 3-stage piston, ac motor driven, 208 v, 190 w, 400 cps, 3 ϕ, 2.41 × 1.524 m diameter, 4.55 kg. Air cooled with own fan.	
	Two stage, cross flow tubing	0.152	0.0381	0.136	Two cycle, four stage, oil activated metallic diaphragm, ac/dc, hydraulic or air turbine driven, 0.208 × 0.216 m diameter.	*j*
15	One stage, coiled tubing	0.0381	0.00955		Oil activated, metallic diaphragm, dc motor driven, 28 v, 1 w, 0.165 × 0.165 × 0.127 m, 4.54 kg.	*j*
	Two stage, coiled tubing	0.0762	0.0127		Oil activated, dual diaphragm, dc motor driven, 28 v, 1 w, 0.178 × 0.178 × 0.127 m, 6.8 kg.	*j*

[g]Filter absorber system is a required accessory. Contact firm for details.
[j]Motor required.

12.2.3. Expansion-Engine Coolers. The expansion-engine cooler is based on the adiabatic and reversible expansion of a gas in doing work on an external load. This expansion results in a reduction of the gas temperature due to the decrease in the internal energy of the gas by the amount of the work done. A reversible expansion of a gas from a high to a low pressure can be produced by either piston- or turbine-type expanders; however, more development work has been done on the piston type because of its ability to handle the small gas flows required. If it is necessary that liquid coolant be produced by an expansion engine, the expanders are used only to produce the low temperature required for liquefaction, never to form the liquid. The formation of liquid in the expander is a highly inefficient process because the liquid wets the wall of the cylinder or turbine, thereby facilitating the flow of heat from the hot walls to the working substance and introducing considerable irreversibility into the cycle. In addition, the liquid introduces serious mechanical instability in the expander and the expander must work much harder to move the vapor-laden gas. The production of the actual liquid is usually carried out by a separate Joule-Thomson expansion valve that is supplied with gas cooled by heat exchange with a regenerative heat exchanger. Two types of expanders are now in use: the piston and the turbine types. Table 12-6 lists specific models of commercially available expansion-engine coolers.

12.2.3.1. *Gifford-McMahon Piston Expander.* The Gifford-McMahon expander uses a reciprocating flow engine based on the Stirling cycle, in which the gas is passed back and forth through a thermal regenerator that acts as a heat reservoir. The Gifford-McMahon expansion engine consists of an expander-regenerator, a compressor, a surge tank, a supply tank, and the necessary controls (Fig. 12-11). The compressor is a special noncontaminating type, usually with a piston or diaphragm arrangement, used to maintain the coolant gas at the necessary pressure. The surge tank and pressure regulators provide a continuous, constant-pressure gas supply for the cooler. The supply tank is provided as a high pressure source of make-up gas to maintain the proper amount of gas in the system for long periods.

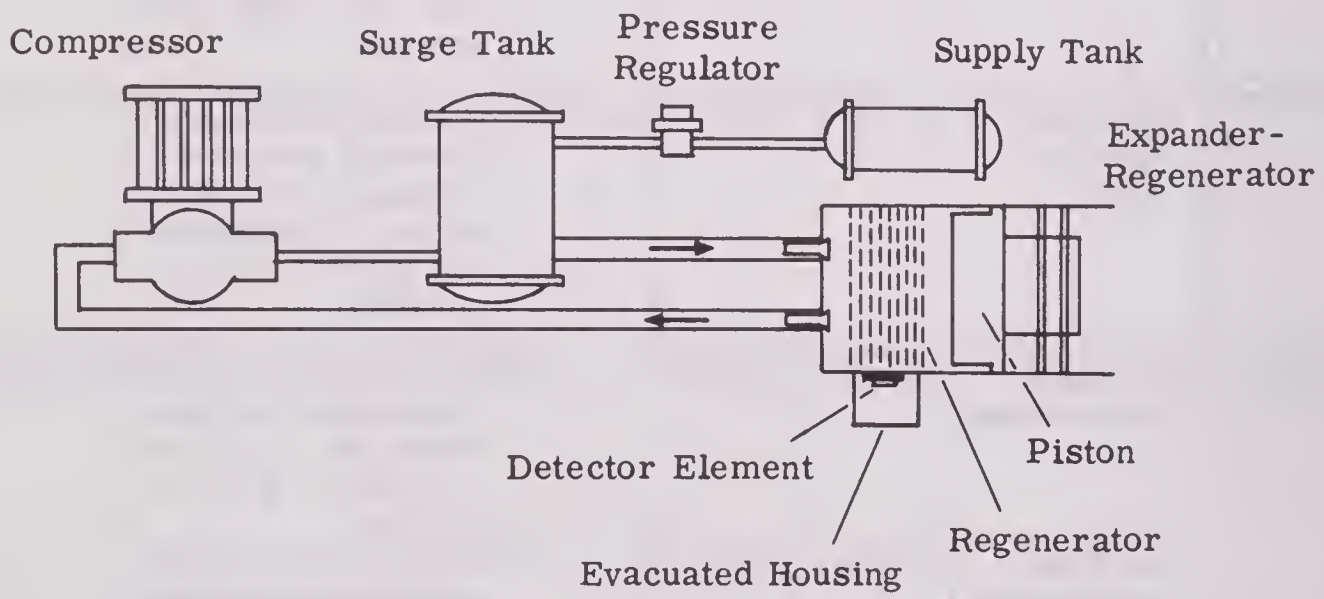


FIG. 12-11. Gifford-McMahon expansion engine.

TABLE 12-6. EXPANSION ENGINE COOLERS

Manufacturer	*Model*	*Cycle*	*Coolant*	*Operating Pressure* [$(kg/m^2) \times 10^3$]	*Flow Rate*
AiResearch	$Ne\text{-}H_2$	Closed, Claude & Joule-Thomson	Ne He	36.7 376	0.36 kg/m 0.015 kg/m
	134708-1-1 N_2	↓	Ne N_2	21.05 49.3	0.036 kg/m 0.08 kg/m
Air Products	Piston Turbine	Closed ↓			
Arthur D. Little	Cryodyne CRYR-200	Closed, Gifford-McMahon & Joule-Thomson	He		
	Minicooler MNR-11	Closed (laboratory use only)	He	176	31.1 liter/min at 60° 19.7 liter/min at 80°
Cryogenerators Div. Norelco	Cryogenerator 42300	Closed, Stirling	He/N_2	84.5	
Malaker Labs[a]	Mark V Mark VII	Closed, Stirling ↓	He	176	
Norden	Integrated Detector Cooler	Closed, Gifford-McMahon	He	210.5	8.5 liter/min
Hughes Aircraft Co.	HAC MK II/12	Solvay engine with compressor	He^4	200/60	0.20 g/min
	HAC MK II/30	Solvay engine with compressor	He^4	200/60	0.17 g/min
	HAC MK III/4	Solvay engine plus Joule-Thomson	He^4	300/150 Solvay 160/10 Joule-Thomson	2.0 g/min 0.2 g/min
	HAC MK IV/80	Stirling refrigerator with air-liquid transfer loop	He^4 + air	150/75 He 10 Air	
	HAC MK IV/30	Stirling refrigerator with neon-liquid transfer loop	He^4 + Ne	150/75 He 10 Ne	

TABLE 12-6. EXPANSION ENGINE COOLERS (*Continued*)

Manufacturer	*Operating Temperature* (°K)	*Cooling Capacity* (w)	*Cooldown Time* (min)		*Expander*	
				Construction	*Size* (m)	*Weight* (kg)
AiResearch	27.16 4.2 ± 0.1 25.66 78	16 1 1 10		Turbine ↓	0.089 × 0.038 diameter	
Air Products				Piston Turbine		
Arthur D. Little	4.3 17 45 60-300	0.25 0.04 0.80 0.05 at 60° 0.17 at 80°	 5-8 3-5	Piston ↓	0.33 × 0.457 × 0.915 0.038 × 0.038 × 0.07 with 0.0794 × 0.0286 diameter motor	61.25
Cryogenerators Div. Norelco	25-300	1 at 30° 10 at 80°	5-7	Displacer	0.1015 × 0.305 × 0.127	4.536- 5.44 with- out motor
Malaker Labs[a]	95 70	1 3	12	Displacer ↓	0.1015 × 0.152 × 0.203 0.356 × 0.127 diameter	
Norden	30-35	0.05	10	Piston	0.178 × 0.0381 diameter	0.51
Hughes Aircraft Co.	10-15	0.4 at 12.0°K	20 min with typical CuGe installation	Piston engine	0.05 diameter × 0.3	1.0
	20-40	2.0 at 30°K 1.0 at 25°K 0 at 20°K	5 min with typical HgGe installation	Piston engine	0.05 diameter × 0.25	1.0
	2.0-4.2	2.0 at 4.2°K +5.0 at 15°K +12.0 at 50°K +60.0 at 150°K	30 min unloaded	Piston engine	0.11 diameter × 0.6	10.0
	70-90	9.0 at 78°K (see Remarks)	<5 min with typical detector flask	Piston		
	28	2.0 at 28°K (see Remarks)	<10 min with typical flask	Piston		

TABLE 12-6. EXPANSION ENGINE COOLERS (*Continued*)

Compressor	*Cooling Head*	*Remarks*
Ne=single stage, oil-lubricated centrifugal, ac motor driven, 200 v, 400 cps, 4.9 hp, 0.152 × 0.152 × 0.279 m. He=two stage, nonlubricated piston, ac motor driven, 200 v, 8 A, 500 cps, 1.75 hp, 0.61 × 0.356 m diameter.	Ne=regenerative heat exchanger He=bundle-type counterflow and plate-fin counterflow	16-w Claude by-pass expander precooler with 1-w Joule-Thomson cooler.
Single stage, nonlubricated piston, ac motor driven, 200 v, 400 cps, 0.42 hp, 0.406 × × 0.203 m diameter. Single stage, nonlubricated piston, ac motor driven, 200 v, 400 cps, 1.0 hp, 0.61 × 0.356 m diameter.	Bundle-type counterflow	Two-stage neon cooler. Two-stage nitrogen cooler.
Single stage, oil activated, nonmetallic diaphragm, motor driven. Single stage, centrifugal compressor, driven by expander.	Bundle-type counterflow	Not now used for detector cooling.
Two stage, oil-lubricated piston, ac motor driven, 440 v, 4 kw, 60 cps, 3 ϕ, plus 110 v, 110 w, 60 cps, 1 ϕ for expander, 0.71 × 0.762 × 1.905 m, 352.5 kg.		Designed for masers, can be modified for detectors.
MNRC-1, single stage, oil-lubricated piston, ac motor driven, 115 v, 350 w, 60 cps, 1 ϕ, plus 110 v, 110 w, 60 cps, 1 ϕ for expander, 0.305 × 0.61 × 0.356 m, 54.5 kg.	Stainless steel cylinder 0.1905 liter × 0.00788 m diameter.	Laboratory use only; requires 10-μ filter.
–	Evacuated container 0.0191 m diameter.	Requires 1750 rpm. motor, power input 200 w at 30°K, 100 w at 100°K.
–	Insulated container 0.0127 m diameter.	Includes ac motor, 110 v, 1.1 A, 125 w, 60 cps, 1 ϕ.
–	Insulated container 0.0127 m diameteer.	Includes ac motor, 208 v, 0.8 A, 140 w, 400 cps, 3 ϕ.
Single stage, nonlubricated diaphragm ac motor driven, 0.305 × 0.089 m diameter, 4.54 kg.	Integral with expander.	Requires surge tank, makeup supply, and pressure regulators.
Single stage 3-piston dry compressor 0.11 m diameter × 0.35 m long, 12.0 kg, 500 w electrical input.	0.0065 m diameter cylinder	All performance figures given for 170°F (350°K), 50,000 ft altitude ambient with self-incorporated heat rejection. Size and weight figures do not include heat-rejection mechanism.
Same as MK II/12 compressor above.	0.015 m diameter cylinder	MK II can be modified for efficient cooling at higher temperature. II/12 and II/30 will refrigerate at any temperature above nominal without modification.
Solvay compressor: single stage 4 piston, 0.3×0.3×0.8 m, 5 kw input, 25 kg. Joule-Thomson compressor: 2 stage 4 pistons, 0.3 × 0.3 × 0.75 m, input 1.5 kw, 25 kg.	0.075 m diameter cylinder.	Not suited for usual infrared systems.
Total system: 0.075 × 0.125 × 0.2 m, 450 w input, 8 kg.	Glass dewar as needed.	Replaces typical LN_2 transfer system; semi-closed air loop uses environment as reservoir with self-filtered intake. Multi-detector systems can be cooled.
Total system: 0.075 × 0.125 × 0.2 m, 450 w input, 8 kg.	Glass dewar as needed.	Similar to MK IV/80 with sealed neon loop. Both MK IV systems are rated with 3 m transfer line length.

[a]Two-stage cooler, consisting of two displacer expanders on a common crankshaft, for operation between 30°K and 90°K under development.

12.2.3.2. *Displacer Piston Expander.* The displacer expander (Fig. 12-12) is a modification of the reciprocating flow engine based on the Stirling cycle, in which the gas is passed back and forth through a thermal regenerator that acts as a heat reservoir. The displacer expansion engine is considerably simpler than a comparable Gifford-McMahon expansion engine in that there are no valves and a separate compressor is not required. The compression of the gas is produced by the motion of the main piston and the displacer piston in coming together. The compressed gas flows through the water jacket, where the heat of compression is removed, to the regenerator. The expansion of the gas takes place against the displacer piston in the space at the top of the cylinder. The main drawbacks of the piston-type expander are the wear of the sliding surfaces and the valve leakage when small volumes of gas are used. In addition, the reciprocating-flow engines are limited by the heat capacity of the regenerator material. Lead is the most common material used, since it can store significant amounts of heat down to approximately 14° K; other materials such as aluminum, zinc, brass, and bronze have no meaningful heat capacity below 35° K.

12.2.3.3. *Turbine Expanders.* Turbine expanders have not been used much because they require relatively larger flows of gas than piston compressors and expanders. Infrared detector coolers requiring low refrigeration loads, and operating at the normal atmospheric boiling point of the gas, usually have flow rates much too low for turbine machinery. Turbine machinery is more applicable to systems using depressed boiling points, where overall life and reliability are important considerations. Progress has been made recently in the development of turbine expanders that operate at low pressure ratios and flow rates, and that are capable of providing sufficient refrigeration to liquefy cryogenic gases. A schematic diagram of a proposed helium refrigerator using a turbine expander for cooling a maser to 4.2° K is shown in Fig. 12-13.

12.2.4. Thermoelectric Coolers [1, 8-12]. Table 12-7 lists specific models of commercially available thermoelectric coolers.

The basic principle of the thermoelectric or Peltier cooler is the Peltier cooling effect, which is caused by the absorption or generation of heat when a current passes through a junction of two dissimilar materials. The rate of pumping of heat is directly proportional to the current, and the constant of proportionality is known as the Peltier coefficient, π. It represents a potential difference, which is determined by the Fermi energy plus a transport energy. Thus, for a current I, the rate of pumping of heat by the Peltier effect, $\dot{Q}_\pi$, is

$$\dot{Q}_\pi = \pi I \tag{12-1}$$

The Seebeck coefficient of a couple, S, is the ratio of the thermal emf to the temperature difference which produces that emf. For a single material, a quantity s, called the Seebeck coefficient of the material, can be defined in such a way that the Seebeck coefficient of any couple is given by the difference of the Seebeck coefficients of the two materials from which it is constructed:

$$S = s_2 - s_1$$

The Seebeck coefficient of a p-type material, s_p, and the Seebeck coefficient of an n-type material, s_n, must have opposite signs. Conventionally, s_p is taken as positive and s_n as negative. Thus, for a couple composed of a p-type and an n-type material, the difference $s_p - s_n$ is actually the sum of two positive numbers. In order to avoid difficulties with signs it is often convenient to use absolute values:

$$|S| = |s_n| + |s_p|$$

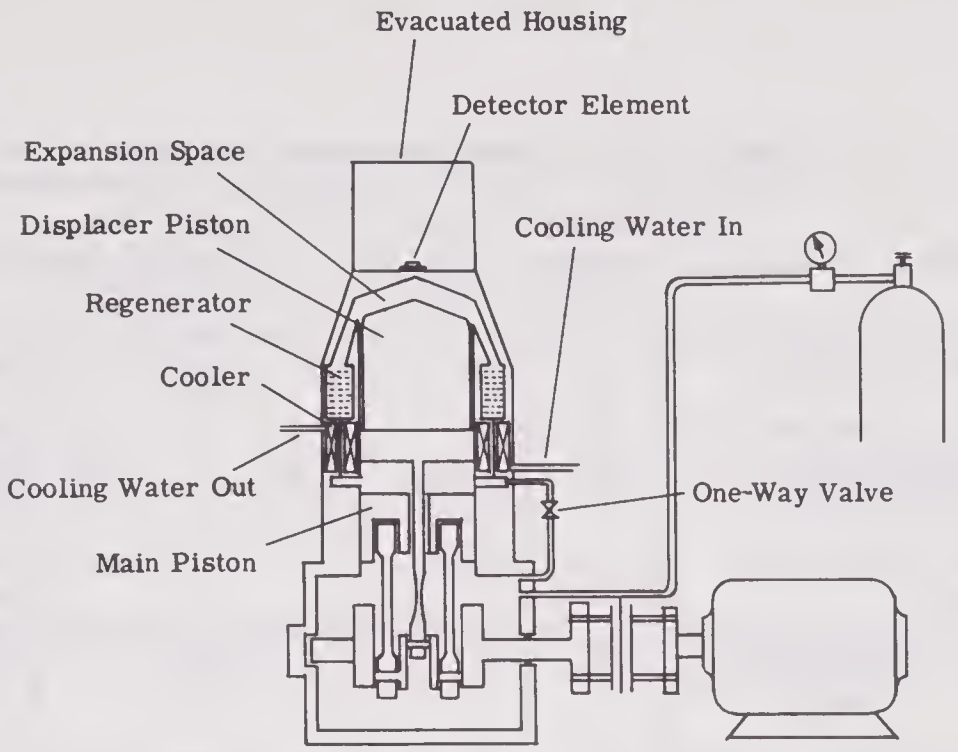
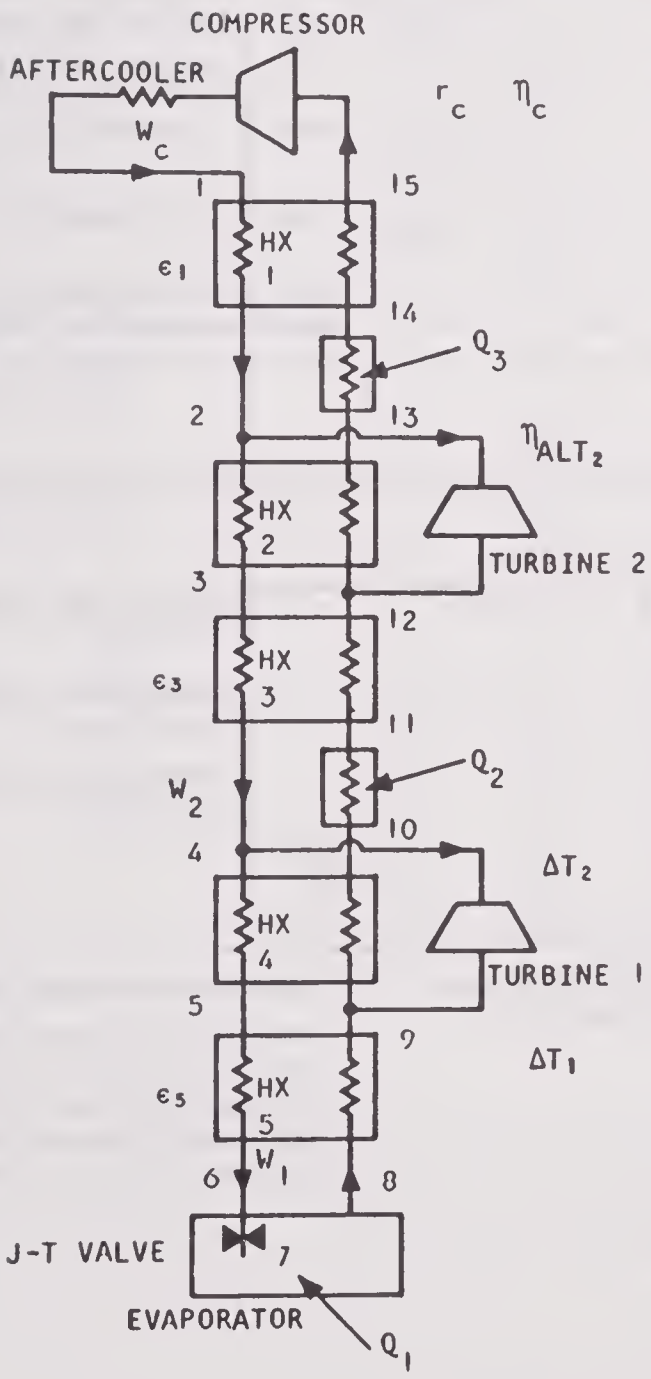


FIG. 12-12. Displacer expansion engine.



MAJOR INPUT PARAMETERS

r_c = Compressor pressure ratio

η_c = Compressor overall efficiency

T_1 = Aftercooler exit temperature

P_8 = Evaporator exit pressure

Q_1 = Primary refrigeration load

Q_2, Q_3 = Secondary cooling loads

$\epsilon_1, \epsilon_3, \epsilon_5$ = Effectiveness of first, third and fifth heat exchangers

$\Delta T_1 = T_9 - T_8$

$\Delta T_2 = T_4 - T_{11}$

$\eta_{ALT_1}, \eta_{ALT_2}$ = Alternator efficiencies

FIG. 12-13. Analytical model for two expander Claude Cycle Analysis Computer Program.

Table 12-7. Thermoelectric Coolers

Manufacturer	*Model*	*Construction*	*Nominal Load Cooling* (mw)	*Cold Junction Temperature* (°K)	*Heat-Sink Load* (w)	*Hot Junction Temperature* (°K)	*Size* (m)
Energy Conversion	D5-60	One-stage	2.6	240		300	0.044 × 0.025 × 0.0094
	D9-60	One-stage	5.3	240		300	0.044 × 0.025 × 0.00686
	E9-60	One-stage	11.2	240		300	0.053 × 0.0348 × 0.00686
	D-series, E-series	Two-stage	*d*	207-225	5-12	300	0.044 × 0.025[c] 0.053 × 0.0348[c]
	D-series, E-series	Three-stage	*d*	200-205	12	300	0.044 × 0.025[c] 0.053 × 0.0348[c]
Materials Electronics (Melcor)		Hot & cold stage with 15 two-stage modules	4	233		293	0.044 × 0.119 o.d.
Pesco	094429	Four-stage cascade	15	195	12	300[a]	0.07625 × 0.117 × 0.152
	094438	Four-stage cascade	15	195	12	300[a]	0.043 × 0.061 × 0.035
	094446	One-stage	50	242 in vacuum 252 in N_2 259 in air	3	300[a]	0.01745 + 0.0254H
	094447	Two-stage cascade	20	223/251	4/4	300/351[a]	0.0413 × 0.0524 × 0.0381
	094492	One-stage, eight couples	200	252	3	300[a]	0.0127 × 0.0088 × 0.0127
	094493	Three-stage cascade	15	208	1	300[a]	0.0305 × 0.0508 × 0.0302
	094567	One-stage	10	243 253 258	0.5	300[a]	0.0159H
	094568	One-stage, three couples	20	243 253 258	2.8	300[a]	0.0206H
	094575	One-stage	30	253	1		
	094618	Two-stage		203	1.9	300[a] 300[a]	0.0381H 0.0305 × 0.0508 × 0.0478
Radiation Electronics	Thermo-electric cooler	Three-stage		195			0.0401H
Texas Instruments	Thermo-electric cooler	One-stage, two couples	80	258	1	300	<0.00254H
Westinghouse	WX-814	One-stage	10 max 5 nom	*b*	36 max	373	0.0396 × 0.0396 × 0.0127 lugs & leads
	WX-816	One-stage	10 max 5 nom	*b*	36 max	373	0.0396 × 0.0396 × 0.0454 lugs & leads
	WX-817	One-stage	1	$\Delta T = 55$			0.01575 × 0.0159 × 0.0127 lugs & leads
	WX-824	Similar to WX-814				523	
	WX-825	Similar to WX-814				623	
	WX-826	Similar to WX-816				523	
	WX-827	Similar to WX-816				623	

[a]Hot-junction temperature given for rating purposes only; all units capable of operating with hot junction temperature of 373°K.

[b]

Westinghouse Model		ΔT
WX 814-E	WX 816-E	35°C
WX 814-F	WX 816-F	40°C
WX 814-G	WX 814-G	45°C
WX 814-H	WX 814-H	50°C
WX 814-J	WX 814-J	55°C

TABLE 12-7. THERMOELECTRIC COOLERS (*Continued*)

Diameter (m)	*Mounting Area*	*Vacuum Pressure* (mm Hg)	*Voltage* (volt)	*Current* (amp)	*Ripple Maximum* (%)	*Weight* (kg)	*Remarks*
–	0.0206 × 0.025	Depends on application	1.2	5	10	0.017	Electrically isolated heat transfer surfaces.
–	0.0206 × 0.025	↓	1.2	9	10	0.014	↓
–	0.03 × 0.0348	↓	2.5	9	10	0.025	↓
–	[c]	↓	[c]	[c]	10	[c]	Long time constant, others available that operate into a low-temperature heat sink.
	[c]	↓					↓
–	[c]	↓	[c]	[c]	10	[c]	↓
0.0525	15 coplanar heat pads	Depends on application	28	1.8	10	0.453	Used to cool vidicon tube; consists of 15 two-stage modules, 8 modules in cold stage, and 20 modules in hot stage.
–	0.0161 m^2	10^{-6}	0.6	20	10	25	Includes integral detector.
0.0356	0.0161 m^2	10^{-6}	0.6	20	10	2.5	
0.025-0.0253	0.0111 m diameter	Depends on application	0.4	7	10	0.5	
0.0325	0.0322 m^2	10^{-3}-10^{-6}	1.2	3	10	2.0	Used for cooling of five diodes.
–	0.0322 m^2	Depends on application	0.85	3.5	10	0.2	
0.0239	0.008 m^2	10^{-6}	0.17	6	10	1.0	
0.00396	0.00396 m diameter	Depends on application	0.1	5	10	0.1	
0.00762	0.00762 m diameter	Depends on application	0.4	7	10	0.2	
0.00762	0.00508 m diameter	Air	0.2	5	10	0.5	Fast time constant, cooldown 5 sec.
0.0239	0.00318 × 0.00635 m	10^{-6}	0.19	10	10	1.0	
0.0302			2.0	2.5			Includes integral detector.
0.0028			0.2	<4	<15		Horizontal cold mounting surface reaches 63% of final ΔT after 1.5 min with no load.
–	0.148 m^2	Air[d]	1.2	20	10	2.8	Horizontal cold mounting surface reaches 63% of final ΔT after 1.5 min with no load.
–	0.103 m^2	Air[d]	1.2	20	10	3.1	Vertical cold mounting surface reaches 63% of final ΔT after 1.5 min with no load
–	0.043 m^2	Air[d]	1.25	6	10	1.5	

[c]Depends on ΔT and cooling load.

[d]Coolers rated with cold junction and device being cooled surrounded by insulation at normal ambient temperature (T_A) of 35°C.

The Seebeck and Peltier coefficients are related by the first of the Thomson relations [13 or 14]:

$$\pi = ST \tag{12-2}$$

It follows that

$$\dot{Q}_\pi = STI \tag{12-3}$$

where T is the temperature of the cold junction when the couple is used for cooling.

Evidently, then, Peltier cooling will be more effective with materials which exhibit large Seebeck coefficients. It is evident also that cooling will be more effective when the joule heating, I^2R, is a minimum. The current, I, cannot be made small without reducing the rate of heat pumping (see Eq. 12-1). Therefore it is necessary to keep the resistance low, which can be accomplished by the use of material with low electrical resistivity, ρ. A third important property is the thermal conductivity κ. Clearly this parameter should be kept as small as possible, since it would be of little value to pump heat from one region to another if most of it could flow back again.

Therefore, the three following parameters serve to characterize a material for its cooling capabilities:

s = Seebeck coefficient (v $°K^{-1}$)

ρ = electrical resistivity (ohm-cm)

κ = thermal conductivity (w cm^{-1} $°K^{-1}$)

These three quantities vary with temperature, and for accurate calculations the variations must be taken into account. Furthermore. variations of s with T give rise to an additional thermoelectric effect, the Thomson effect, and a rigorous treatment must take into account the Thomson heat which arises when a current and a parallel temperature gradient exist in a material for which s varies with T. However, for the purposes of this discussion of Peltier cooling, s, ρ, and κ will be assumed constant over the temperature range under consideration, and hence the Thomson effect will not affect the problem.

12.2.4.1. *Figure of Merit.* The three material properties s, ρ, and κ frequently appear in the same combination in discussions of thermoelectronic devices. This combination is usually referred to as the figure of merit of the material and is denoted by Z, where

$$Z = \frac{s^2}{\rho\kappa} \ (°K^{-1}) \tag{12-4}$$

In general, the larger the figure of merit, the more useful the material for thermoelectric applications. At present, the best values of Z for semiconducting materials are slightly greater than $3 \times 10^{-3}/°K$. (The best values for metals are about $0.1 \times 10^{-3}/°K$.) A figure of merit of $3 \times 10^{-3}/°K$ makes it possible, for instance, to pump heat from ice to steam with a single-stage device.

The figure of merit of a couple is defined by

$$Z_c = \frac{S^2}{RK} \tag{12-5}$$

where the material constants ρ and κ of Eq. (12-4) have been replaced by the electrical resistance R and the thermal conductance K.

The maximum possible value of Z_c for any couple composed of two given materials ([14], page 11) is given by

$$Z_{max} = \left(\frac{S}{\sqrt{\rho_1\kappa_1} + \sqrt{\rho_2\kappa_2}} \right)^2 \tag{12-6}$$

It should be noted that Z_{max} depends only on the properties of the materials, whereas Z_c depends also on the dimensions of the couple. Z_c approaches Z_{max} only under ideal conditions when the relative values of the various dimensions have been properly adjusted.

If the values of ρ and κ are the same for the two arms of the couple, it can easily be seen that $Z_c = Z_{max}$ when the dimensions of the two arms are equal. If, in addition, the couple is made from p-type and n-type materials such that $s_p = -s_n \equiv s$, then $Z_c = Z_{max} = s^2/\rho\kappa$.

12.2.4.2. *Performance of a Peltier Couple.* This treatment is not intended to be an exact physical or mathematical description, but only an outline of the important features of Peltier couples. The numerical examples given are evaluated from approximate expressions because the errors are small and the properties of the thermoelectric materials are not known with sufficient accuracy to warrant exact calculations.

It will be assumed that the couple has Seebeck coefficients s_n and s_p, electrical resistivities ρ_n and ρ_p, thermal conductivities κ_n and κ_p, total series electrical resistance R, and total parallel thermal conductance K between the ends, which are at temperatures T_c and T_h. The temperature difference $T_h - T_c$ will be called ΔT. Figure 12-14 is a sketch of the idealized Peltier couple that will be discussed. The conductor connecting the thermoelectric materials is assumed to have zero Seebeck coefficient, zero electrical resistivity, and infinite thermal conductivity. The thermal and electrical resistances at these connections are assumed to be zero. When it is also assumed that the properties of the material are independent of temperature, it is possible to give a complete description of the equilibrium performance of the couple. It will be shown that s_n, s_p, R, K, and T_c completely define the maximum temperature difference attainable with this particular couple. However, this is not necessarily the maximum temperature difference attainable with the materials from which the couple is constructed.

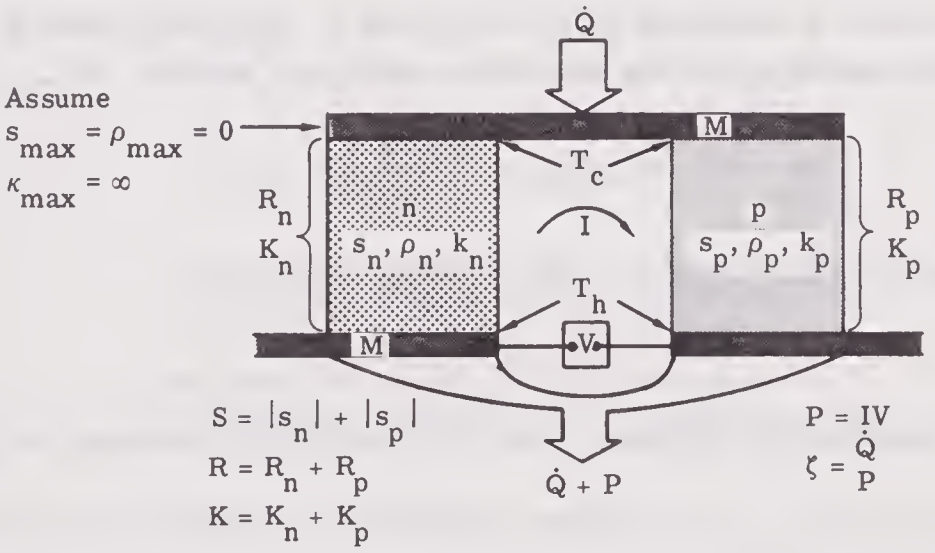


FIG. 12-14. Peltier couple with notation and definitions.

Four quantities are of interest: the pumping current I, the heat-pumping rate Q, the temperature difference ΔT, and the coefficient of performance ζ (to be defined later). Only two of these are independent. However, generally two conditions are of greatest interest, namely, the pumping of the maximum amount of heat, and the most efficient pumping of heat. Each of these conditions fixes the magnitude of the current so that, for a given couple, the heat-pumping capacity and coefficient of performance can be represented as functions of ΔT and T_c only.

12.2.4.3. *Heat Balance at the Cold Junction.* In the couple shown in Fig. 12-14, the current passing from n to M to p will remove heat from M. The rate of heat removal by the Peltier effect is $\dot{Q}_\pi = (s_p T_c I - s_n T_c I) = S T_c I$.

It can also be shown that the uniform generation of joule heat throughout the thermoelectric material results in a flow of heat to each of the junctions at the rate $\dot{Q}_j = \frac{1}{2} I^2 R$. (It is assumed that there is no heat transfer through the sides of the thermoelectric arms.) Finally, the heat conducted back to the cold junction from the hot junction is $\dot{Q}_f = K \Delta T$.

At equilibrium, the difference between the flow of Peltier heat, $\dot{Q}_\pi$, from the cold junction and the flow of joule heat, $\dot{Q}_j$, and conducted heat, $\dot{Q}_f$, to the cold junction represents the rate $\dot{Q}$ at which the couple pumps heat. Thus,

$$\dot{Q} = \dot{Q}_\pi - (\dot{Q}_j + \dot{Q}_f) = S T_c I - \tfrac{1}{2} I^2 R - K \Delta T \tag{12-7}$$

Because of the power dissipated in the thermocouple (P) during the heat-pumping process, $\dot{Q} + P$ watts must be rejected at the hot junction for $\dot{Q}$ watts absorbed at the cold junction.

12.2.4.4. *Maximum Rate of Heat Pumping.* The current which produces the maximum rate of heat pumping is easily obtained maximizing $\dot{Q}$ (by setting $\partial \dot{Q} / \partial I = 0$). The result is that the current for maximum steady-state heat pumping, $I_{\dot{Q}}$, is

$$I_{\dot{Q}} = S T_c / R \tag{12-8}$$

which is proportional to the temperature of the cold junction and independent of the load or temperature difference. The maximum rate of heat pumping is obtained by substituting Eq. (12-8) into Eq. (12-7). The result is

$$\dot{Q}_{max} = \frac{1}{2} \frac{S^2 T_c^2}{R} - K \Delta T \tag{12-9}$$

If the cold junction is insulated so that no heat is absorbed from its surroundings, ΔT will rise to the maximum value that this couple can provide, ΔT_{max}. Thus,

$$\Delta T_{max} = \frac{1}{2} \frac{S^2 T_c^2}{RK} \tag{12-10}$$

The figure of merit of the couple, Z_c, may then be defined as

$$Z_c = S^2 / RK \tag{12-11}$$

The maximum temperature difference that this couple can maintain is therefore

$$\Delta T_{max} = \tfrac{1}{2} Z_c T_c^2 \tag{12-12}$$

With optimum design, a couple made of the same materials can maintain a temperature difference ΔT_M, given by

$$\Delta T_m = \tfrac{1}{2} Z_{max} T_c^2 \tag{12-13}$$

where Z_{max} is defined by Eq. (12-6). In general, ΔT_{max} is slightly less than ΔT_M.

Equations (12-9) and (12-10) show that the rate of heat pumping may be written

$$\dot{Q}_{max} = K(\Delta T_{max} - \Delta T) = K \Delta T_{max} \left(1 - \frac{\Delta T}{\Delta T_{max}}\right) \tag{12-14}$$

which decreases linearly from $K \Delta T_{max}$ at $\Delta T = 0$ to 0 at $\Delta T = \Delta T_{max}$.

The coefficient of performance, ζ, is defined as

$$\zeta = \dot{Q}/P \tag{12-15}$$

where P is the power required to pump heat at the rate $\dot{Q}$, and is given by

$$P = IS\,\Delta T + I^2R \tag{12-16}$$

$$= 2K\,\Delta T_{max}\left(1 + \frac{\Delta T}{T_c}\right) \tag{12-16'}$$

By combining Eq. (12-14), (12-15), and (12-16′), one finds the coefficient of performance for maximum heat pumping to be

$$\zeta_{\dot{Q}} = \frac{1}{2}\left(1 - \frac{\Delta T}{\Delta T_{max}}\right)\left(1 - \frac{\Delta T}{T_h}\right) \tag{12-17}$$

This has a maximum value of 1/2 and decreases to zero when $\Delta T = \Delta T_{max}$.

Summarizing; it is now possible to obtain:

(a) The maximum no-load temperature difference, ΔT_{max}, from Eq. (12-12).

(b) The current required for maximum cooling, $I_{\dot{Q}}$, from Eq. (12-8).

(c) The amount of heat pumped, $\dot{Q}_{max}$, for a given ΔT, from Eq. (12-14).

(d) The coefficient of performance, $\zeta_{\dot{Q}}$, for a given ΔT, from Eq. (12-17).

"Typical values of $\dot{Q}_{max}$, $\zeta_{\dot{Q}}$, and $I_{\dot{Q}}$, for material constants somewhat less than the best attainable are presented in Fig. 12-15.

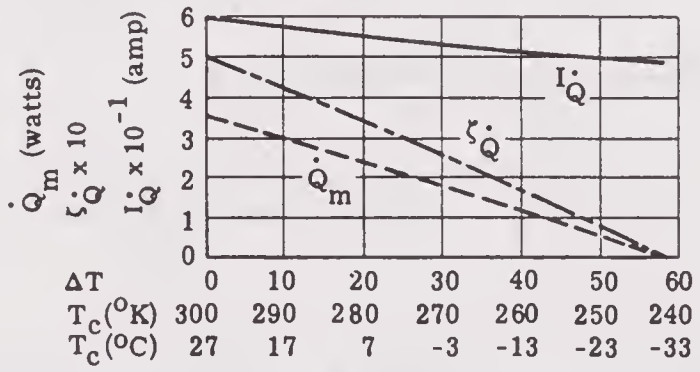


FIG. 12-15. Maximum rate of heat pumping, current required for maximum heat-pumping capacity, and coefficient of performance for maximum heat pumping as a function of T_c (or ΔT) for $T_h = 300°$ K [8].

12.2.4.5. *Maximum Efficiency.* The expression for the current required for maximum efficiency may be obtained by maximizing the coefficient of performance with respect to current, using Eq. (12-7), (12-15), and (12-16). The result obtained shows that the current giving maximum coefficient of performance, I_ζ, is

$$I_\zeta = I_{\dot{Q}}\,\frac{\Delta T}{\Delta T_{max}}\,A \tag{12-18}$$

where A is a factor close to unity given by the relation

$$A = \frac{T_c + \sqrt{T_c^2 + 2T_c\,\Delta T_{max} + \Delta T\,\Delta T_{max}}}{2T_c + \Delta T} \cong 1 + \frac{T_{max} - \Delta T}{2T_c} \tag{12-19}$$

Therefore,

$$I_\zeta \cong I_{\dot{Q}} \frac{\Delta T}{\Delta T_{max}} \tag{12-20}$$

Note that for constant T_c the current for maximum heat pumping, $I_{\dot{Q}}$, is constant and the current for maximum efficiency, I_ζ, varies almost linearly with ΔT.

From Eq. (12-7) and (12-8), the rate of heat pumping is found to be

$$\dot{Q}_\zeta = K\,\Delta T \left(2A - \frac{\Delta T}{\Delta T_{max}} A^2 - 1\right) \tag{12-21}$$

where A is defined by Eq. (12-19) and is nearly unity, so that

$$\dot{Q}_\zeta \cong K\,\Delta T \left(1 - \frac{\Delta T}{\Delta T_{max}}\right) \tag{12-22}$$

This may be written in terms of the maximum heat pumping rate as

$$\dot{Q}_\zeta \cong \dot{Q}_{max} \frac{\Delta T}{\Delta T_{max}} \tag{12-23}$$

and will be zero for $\Delta T = 0$ and $\Delta T = \Delta T_{max}$, with a maximum at $\Delta T = \Delta T_{max}/2$, where $\dot{Q}_\zeta = \dot{Q}_{max}/2$.

The maximum coefficient of performance, from Eq. (12-21), (12-18), (12-16), and (12-15) becomes

$$\zeta_{max} = \frac{1}{2}\left(2 - \frac{1}{A} - \frac{A\,\Delta T}{\Delta T_{max}}\right)\left(\frac{\Delta T}{T_c} + \frac{A\,\Delta T}{\Delta T_{max}}\right)^{-1} \tag{12-24}$$

and again, since A is nearly unity, this simplifies to

$$\zeta_{max} \cong \frac{1}{2}\left(1 - \frac{\Delta T}{\Delta T_{max}}\right)\left(\frac{\Delta T}{\Delta T_c} + \frac{\Delta T}{\Delta T_{max}}\right)^{-1} \tag{12-25}$$

This may be written

$$\zeta_{max} \cong \frac{\Delta T_{max}}{\Delta T} \cdot \frac{1}{2}\left(1 - \frac{\Delta T}{\Delta T_{max}}\right)\left(1 - \frac{\Delta T}{T_h}\right)\left(1 + \frac{\Delta T_{max} - \Delta T}{T_h}\right)^{-1} \tag{12-26}$$

which, from Eq. (12-17), is very nearly

$$\zeta_{max} \cong \frac{\Delta T_{max}}{\Delta T} \cdot \zeta_{\dot{Q}} \tag{12-27}$$

Note that two approximations are used in arriving at this simple form. The exact expression given by Eq. (12-24) can be approximated by Eq. (12-25), since A is very nearly unity. Equation (12-26) is merely Eq. (12-25) written in a different form. Finally, Eq. (12-26) is written as Eq. (12-27) by assuming that

$$\Delta T_{max} - \Delta T << T_h$$

The ratio $\Delta T/\Delta T_{max}$ is seen to be very nearly the ratio of the coefficients of performance for maximum cooling rate and maximum efficiency. Figure 12-16 shows the values of ζ_{max}, I_ζ, and $\dot{Q}_\zeta$ when the couple used for Fig. 12-15 is used under conditions of maximum coefficient of performance. The curves were obtained using Eq. (12-20), (12-23), and (12-27). For very small ΔT, Eq. (12-26) was used instead of Eq. (12-27).

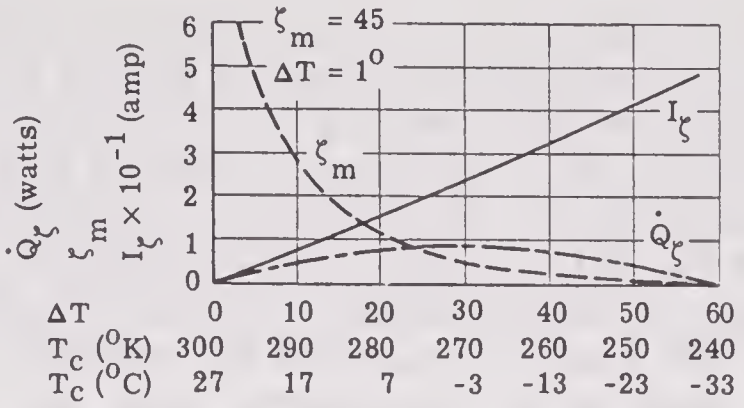


FIG. 12-16. Maximum coefficient of performance, current required for maximum coefficient of performance, and heat-pumping rate for maximum coefficient of performance for typical couple [8]. $T_h = 300°$ K.

Analytical expressions for $\dot{Q}$ and ζ are not difficult to obtain in simple form for other values of current. Some idea of the way in which $\dot{Q}$ and ζ vary with current and temperature difference may be obtained from Fig. 12-17, where $\dot{Q}$ and ζ are plotted against I with ΔT as a parameter. Again, the numbers refer to the couples used for Fig. 12-15 and 12-16.

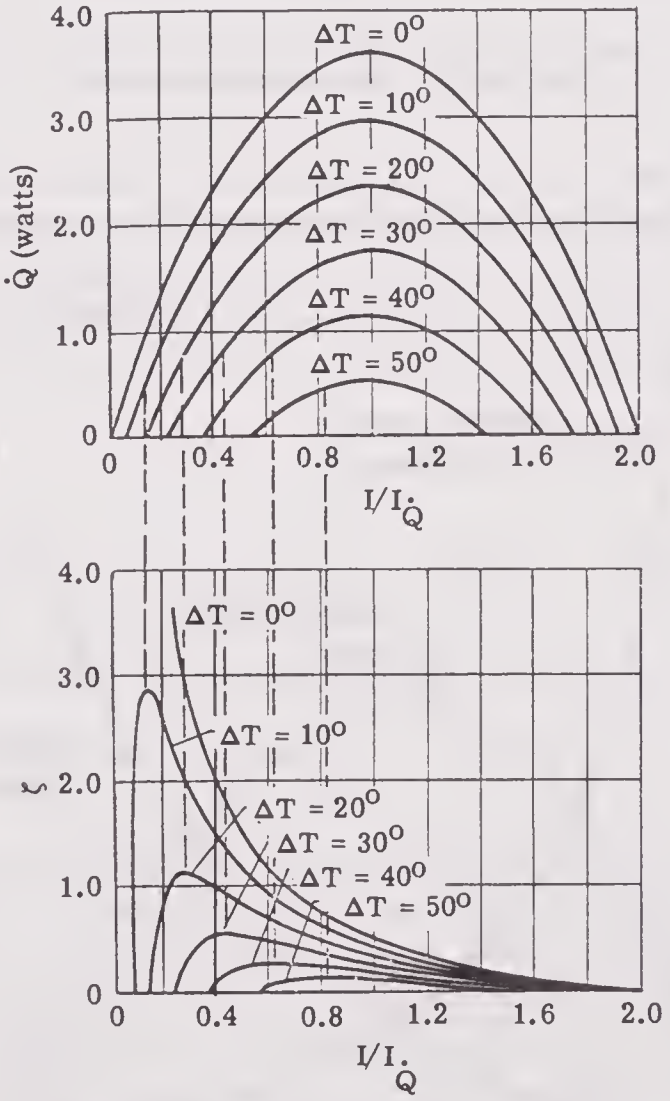


FIG. 12-17. Rate of heat pumping and coefficient of performance as functions of current for various temperature differences with $T_h = 300°$ K [8].

12.2.4.6. *Single-Stage Coolers* [8]. A single-stage thermoelectric cooler consists of a *p*-type and an *n*-type semiconductor connected together by a metallic conductor (Fig. 12-18). An external battery causes the flow of a current; flow produces a temperature difference between the two junctions by absorbing heat at one and releasing it at the other. The performance of a thermoelectric couple as a function of temperature difference is shown in Fig. 12-19. The hot junction is held fixed, and the temperature is varied by changing it at the cold junction. For each change in temperature, the applied voltage has been adjusted to maximize the coefficient of performance. The two main operating modes for any thermoelectric cooler occur at the points where both the heat-pumping rate and the coefficient of performance are maximum.

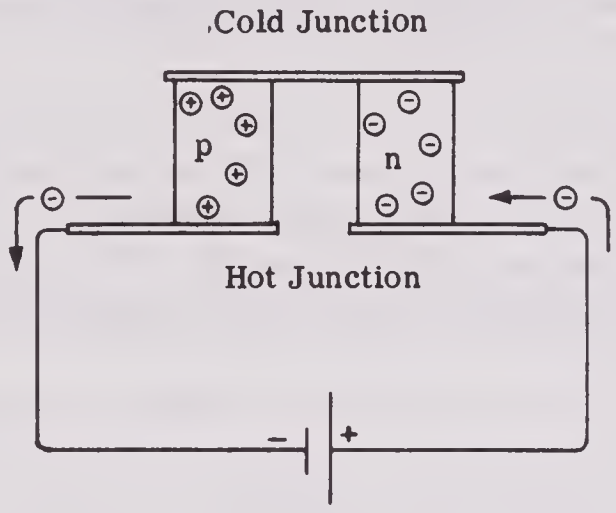


FIG. 12-18. Thermoelectric cooler.

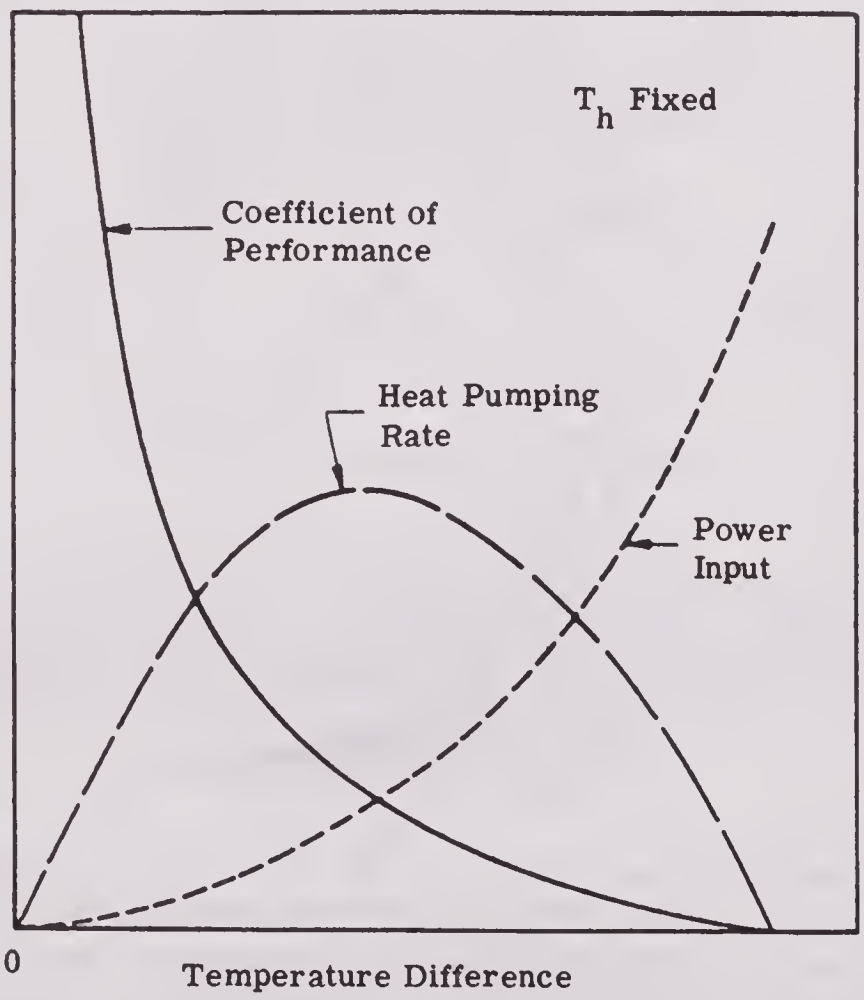


FIG. 12-19. Performance of thermoelectric couple as a function of temperature difference [9].

12.2.4.7. *Single-Stage, Multicouple Coolers.* Thermoelectric couples of similar material can be arranged in thermal parallel to increase their heat-pumping capacity. When identical couples are placed in thermal parallel (Fig. 12-20) and supplied with equal currents, they pump as many times more heat at the same temperature difference and with the same coefficient of performance as there are couples. To ensure that the couples are supplied with equal currents, they are connected in series electrically. Thin foils or films, or materials of good thermal conductivity, are used as insulation between couples since they do not increase the thermal resistance significantly.

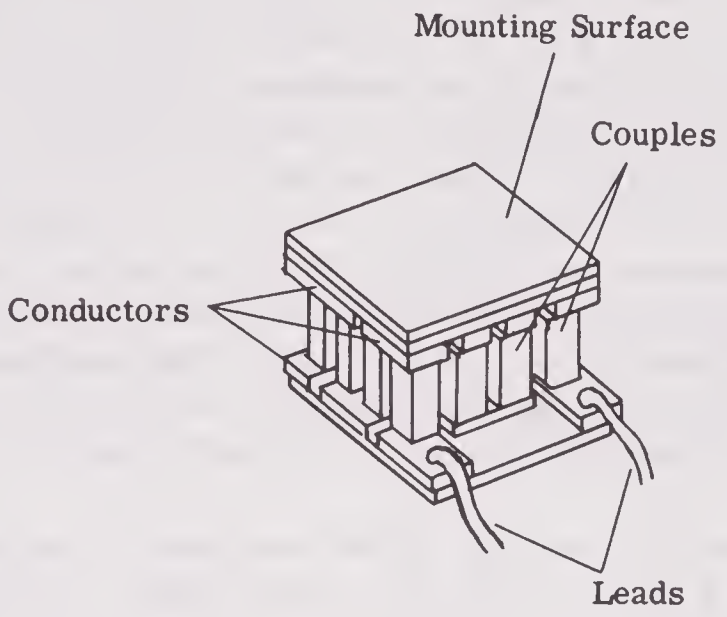


FIG. 12-20. Eight couple, single-stage cooler.

12.2.4.8. *Multistage Coolers.* Thermoelectric couples are connected in thermal series or cascaded for two purposes: to provide a temperature difference greater than that obtainable from a single couple, or to achieve a higher coefficient of performance for a given heat load at an established temperature difference. The limit on the coefficient of performance is related to the limit on the maximum temperature difference by

$$\zeta_Q = \frac{1}{2}\left(1 - \frac{\Delta T}{\Delta T_{max}}\right)\left(1 - \frac{\Delta T}{T_h}\right) \tag{12-28}$$

$$\zeta_{max} \cong \frac{\Delta T_{max}}{\Delta T} \cdot \frac{1}{2}\left(1 - \frac{\Delta T}{\Delta T_{max}}\right)\left(1 - \frac{\Delta T}{T_h}\right)\left(1 + \frac{\Delta T_{max} - \Delta T}{T_h}\right)^{-1} \tag{12-29}$$

The coefficient of performance for cascaded couples is given by

$$\zeta = \left[\left(1 + \frac{1}{\zeta_1}\right)\left(1 + \frac{1}{\zeta_2}\right) \cdots \left(1 + \frac{1}{\zeta_n}\right) - 1\right]^{-1} \tag{12-30}$$

The results obtained from the use of a two-stage cooler may be illustrated by the the following example. A single-stage cooler made from materials such that $Z_c = 2 \times 10^{-3}/°K$ (Sec. 12.2.4.1) and operating between 260° K and 300° K is replaced by two stages, each operating with $\Delta T = 20°$. When the currents are adjusted for maximum efficiency, the coefficients of performance are 0.30 for the single-stage cooler and 0.42 for the two-stage cooler. However, if the currents are adjusted for maximum heat pumping, the coefficients are 0.18 for the single-stage cooler and only 0.06 for the two-stage cooler.

Further improvement could be obtained by adding a third stage, but for the above example the improvement is small (about 10%). In general, the improvement for

two stages is greatest when $\Delta T/\Delta T_{max}$ is nearly unity. When $\Delta T/\Delta T_{max} > 1$, a multistage device is the only way to remove heat from the cold junction under steady-state operation.

The preceding discussion has referred primarily to cascaded systems in which the objective is to improve efficiency with a given temperature difference. If the principal objective is to obtain a large temperature difference between the cooled region and the heat sink, each stage can be operated, inefficiently, at nearly its maximum temperature difference, thereby giving $\Delta T_{max} = \Delta T_{max_1} + \Delta T_{max_2} + \Delta T_{max_3} + \ldots + \Delta T_{max_n}$ for n stages. Since operation at with $\Delta T = \Delta T_{max}$ is inefficient, successive stages at higher temperatures must be designed to pump rapidly increasing amounts of heat. It should be borne in mind that ΔT_{max} is different for each of the successive stages even if constant material properties are assumed, because of the dependence of ΔT_{max} on the square of the absolute temperature of the cold junction.

In considering the design of multistage coolers, it is clear that it would be undesirable to supply power to each stage separately through copper leads from a power supply at a temperature of approximately 300° K. Both the thermal losses through the leads and the thermal insulation between stages, introduced by the required electrical insulation, would constitute sources of inefficiency. This inefficiency can be avoided by designing a network of unequal thermoelectric elements that would allow the optimum currents to be supplied to each aim of each stage from the adjacent stages (Fig. 12-21).

Figure 12-22 shows the variation of the cold-junction temperature as the input current is varied for a typical four-stage cascade cooler. The cold-junction temperature *versus* cold-junction load is shown in Fig. 12-23, and the cold-junction temperature *versus* time is shown in Fig. 12-24. Because the cold-junction temperature is extremely sensitive to convection loading, the thermoelectric cooler is usually mounted in a chamber evacuated to a very low pressure. Figure 12-25 shows the effect of the vacuum pressure on the cold-junction temperature.

For the most effective cooling, the difference in the Seebeck coefficients should be as great as possible, the electrical resistance low, and the thermal conductivity small. In addition, consideration must be given to the optimum cross sections of the arms of

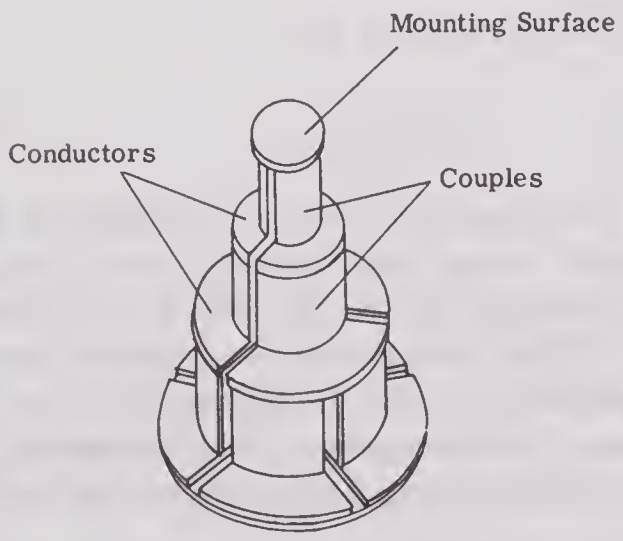


FIG. 12-21. Ten couple, three-stage cascade cooler.

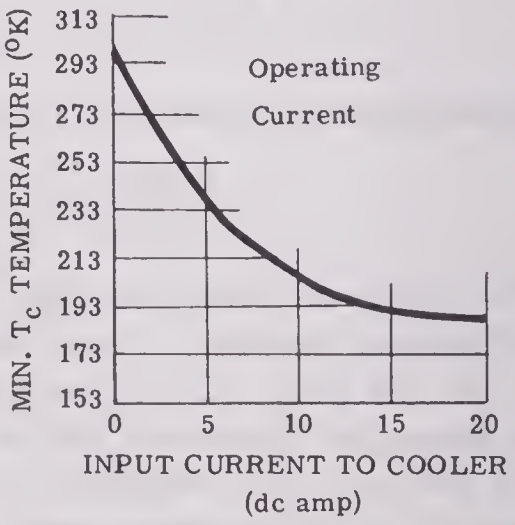


FIG. 12-22. Cold junction temperature *versus* input current with hot junction at +27° C and no thermal load [11]. (*Copyright Pesco Products.*)

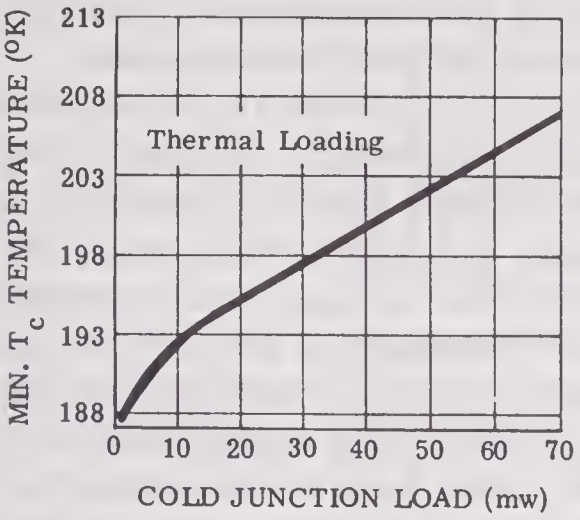


FIG. 12-23. Cold junction temperature *versus* cold junction load with hot junction at +27°C [11]. (*Copyright Pesco Products.*)

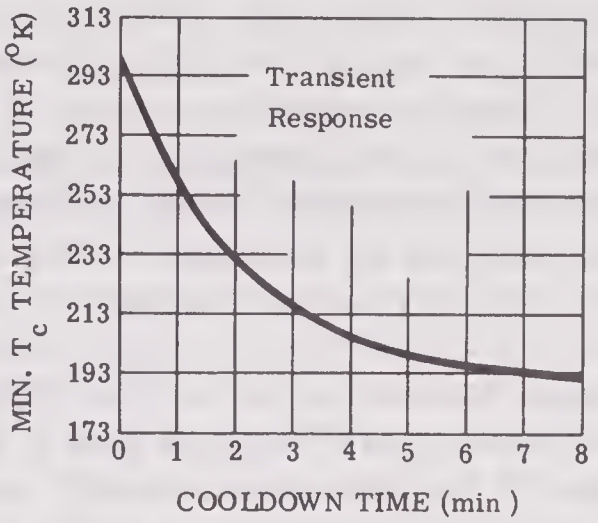


FIG. 12-24. Cold junction temperature *versus* time after optimum current is applied [11]. (*Copyright Pesco Products.*)

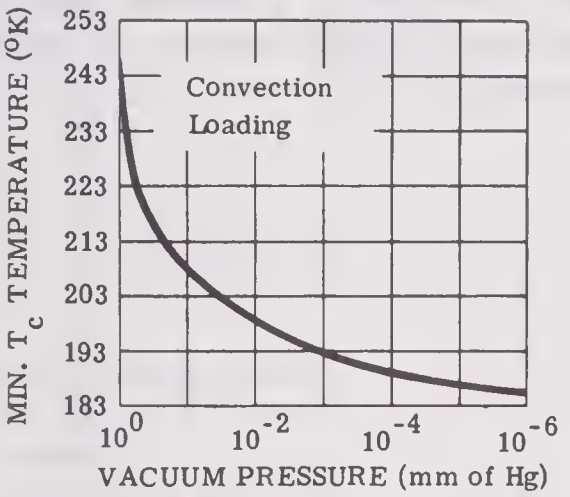


FIG. 12-25. Cold junction temperature *versus* vacuum pressure with hot junction at +27°C and no other thermal load [11]. (*Copyright Pesco Products.*)

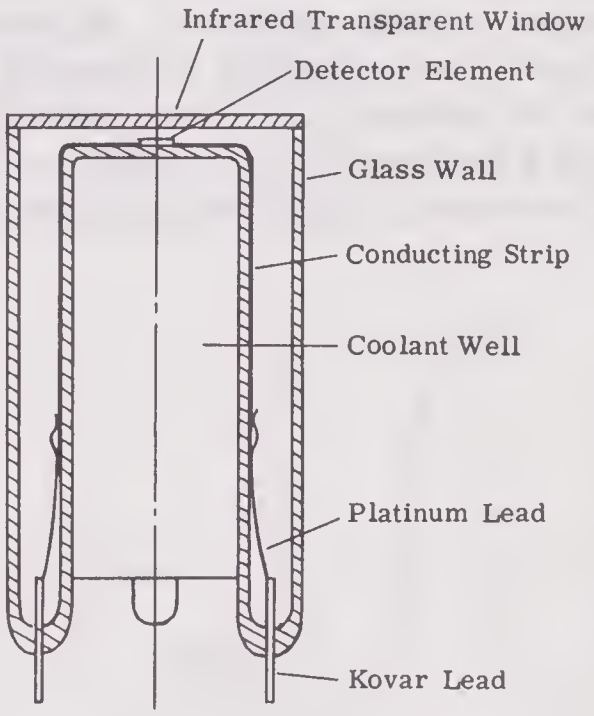


FIG. 12-26. Construction of a single-cell detector dewar.

the couples, heat exchange at the hot and cold junctions, electrical and thermal resistance of the junctions and conductors, power-supply problems, and changes in the heat load or supply current. The design of thermoelectric coolers is given in [8] and [12].

12.2.5. Dewars. Most detectors that require cooling are mounted in a double-walled vacuum enclosure called a dewar. Dewars can be conveniently classified according to the number of coolant cells: single cell (Fig. 12-26) or double cell (Fig. 12-27). The dewar maintains the coolant in thermal contact with the detector, insulates the coolant from the environment to prevent rapid evaporation, prevents frost from collecting over the external surfaces of the dewar (in particular the window), and protects the delicate detector. Also, the vacuum construction protects the detector from the deteriorating effects of the atmosphere.

12.2.5.1. *Dewar Construction.* A variety of detector configurations and envelope constructions are available depending on the mode and temperature of operation. Dewars are usually custom designed for a particular application; however, certain stock dewars

are available (Table 12-8). Dewars are also made by the manufacturers of detectors for use with their detector cells, although some companies will sell them separately. Currently, dewars are constructed mainly of Pyrex glass, although metal is sometimes used. Pyrex glass has an advantage in that it is readily available in many sizes, but for dewars are constructed mainly of Pyrex glass, although metal is sometimes used. Pyrex glass has an advantage in that it is readily available in many sizes, but for helium systems it has one drawback; it is permeable to helium gas at room temperature. Thus, after several uses, the helium dewar must be repumped to eliminate the gas in the space between the walls. This difficulty can be avoided if care is taken never to fill the helium dewar with gas until it is at the temperature of liquid nitrogen. The surfaces of the dewars are normally coated with a thin layer of a substance having good reflection characteristics, except for a thin window or slit on opposite sides of the dewar walls for observation of the liquid level. In the end-looking detector, an infrared-transparent window is joined to the end of the dewar opposite the detector. In the side-looking detector, the window is incorporated in the outside wall of the dewar opposite the detector. Once the window is attached, the detector surface is no longer available for testing. A demountable dewar (Fig. 12-28) that is evacuated before each test is used if it is necessary that the detector surface be readily available. When the detector cannot be physically located in the dewar, a thermal conductor such as a synthetic ruby or sapphire rod provides thermal contact between the detector and the coolant in the dewar to assure efficient cooling.

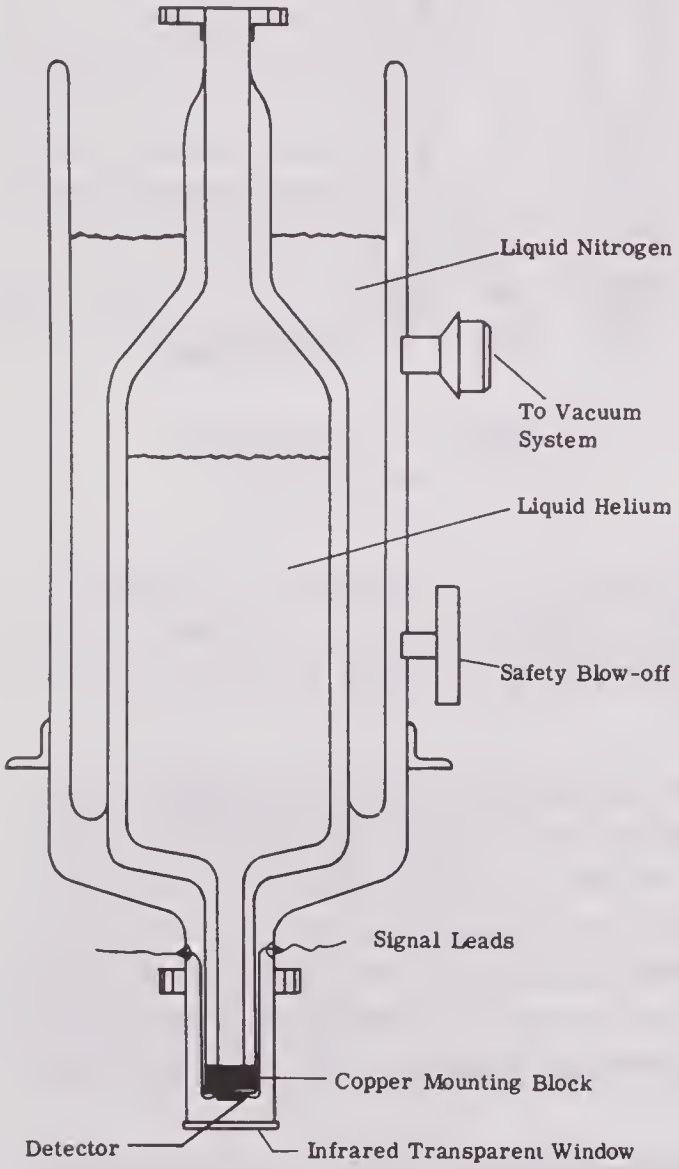


FIG. 12-27. Construction of a double-cell detector dewar.

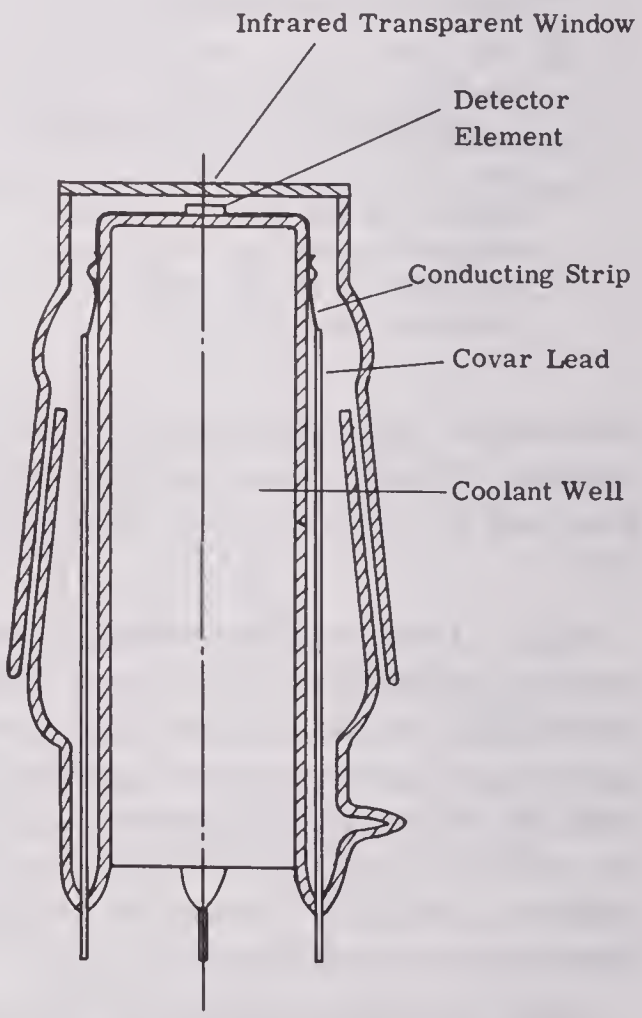


FIG. 12-28. Construction of a demountable dewar.

Table 12-8. Commercially Available Dewars.*

Model/ Dwg No.	Type	Liquid	Capacity (liter)	Operating Time (hr)	Outside Diameter (m)	Well Diameter (m)	Length (m)	Weight		Aperture Size	Detector Size (m)	Operating Attitude	Window
								Empty (kg)	Full (kg)				
C-1322	Direct contact, copper mounting block	N_2	1.7	5-6	0.127[a]	0.1015	0.457	2.27	3.18	0.00435 m², conical spread, variable to fit customer specifications	1×10^{-6} to 5×10^{-6} cube	Vertical with window facing down	Supplied by customer; holder is removable for interchange of windows and filters
C-1595	↓	He	0.7	3-4	0.1049[a]	0.0855	0.555	2.27	2.5	↓	↓	↓	↓
C-2432		N_2	1.25	8-10	0.152[b]	0.0855	0.428	3.52	4.53				
		He	1.0										
C-1901		N_2	2.5	12-14	0.203[b]	0.114	0.482	7.25	9.55	0.03 m diameter	0.02-0.03		
		He	2.0							↓	↓		
D-2064		He	1.5	6-8	0.152[b]	0.122	0.487	2.945	3.18				
D-1416		N_2	1.25	8-10	0.152[b]	0.0855	0.52	4.08	4.95	0.0127 m diameter min at 180° typical	–		
		He	1.0										
D-2061	Direct contact, copper conducting shields	N_2	1.0	3-4	0.152[b]	0.0855	0.487	4.95	5.45	0.0508 m diameter; others available	–	Any position	
		He	0.5									↓	
D-2063	↓	N_2	3.0	10-12	0.152[b]	0.133	0.487	3.18	5.45	0.0222 m diameter, others available	–		
D-1999		N_2	0.5	2-3	0.0865[a]	0.0762	0.432	2.04	2.27	0.0635 m diameter, others available			

*Manufactured by Hofman Laboratories.
[a]Plus space for vacuum connections.
[b]Plus mounting ring and vacuum connections.

Dewars can be used interchangeably with different coolants, but when liquid helium is used, certain precautions must be followed. The liquid must be protected from the atmosphere at all times, since at the liquid-helium temperature all other substances are solids. Therefore a high-capacity vacuum pump is required. Because of the low latent heat of liquid helium, all heat leaks into the bath must be reduced to minimum. This entails the use of a double-cell dewar whose outer container or shield contains liquid nitrogen so that the liquid helium is surrounded by a body at a temperature well below ambient room temperature. The dewar must be cooled down to the liquid-nitrogen temperature before liquid helium is placed in the inner cell to prevent the helium gas from permeating into the interspace and to remove the sensible heat of the dewar with economical liquid nitrogen rather than expensive liquid helium. Because of the low latent heat of helium, a very large quantity would be required to cool the dewar from room temperature to liquid-helium temperature. Whereas in a regular dewar the interspace is evacuated to a low pressure, usually 10^{-6} mm Hg, in a helium dewar a considerable quantity of air, approximately 5 mm Hg, is left in this space. The air acts as a heat conductor and is useful in cooling the well to the liquid-nitrogen temperature, thus reducing the heat load on the liquid helium. When the liquid helium is added to the well, the air in the interspace freezes and provides an excellent insulating vacuum of about 10^{-10} mm Hg.

Temperatures below the boiling point of liquid helium can be produced by controlling the vapor pressure through pumping. A peculiar creep property of liquid helium below 2.19°K permits it to flow against gravity to regions that are warmer. This superfluid helium is called helium II. To reduce the flow of this liquid, a constriction is placed at the top of the cooling well. However, because of the zero viscosity of liquid helium II, a large quantity still passes through the aperture. Large vacuum pumps capable of pumping 2500 liters/min at a pressure of 0.1 mm Hg are required to remove the evaporated gas.

12.2.5.2. *Special-Construction Dewars.* In addition to the cooled-detector-type dewar, special dewars have been built that have cooled enclosures and apertures. These dewars are the same as an ordinary dewar, except the detector is mounted inside the coolant well (Fig. 12-29). The aperture well is coated with an opaque substance

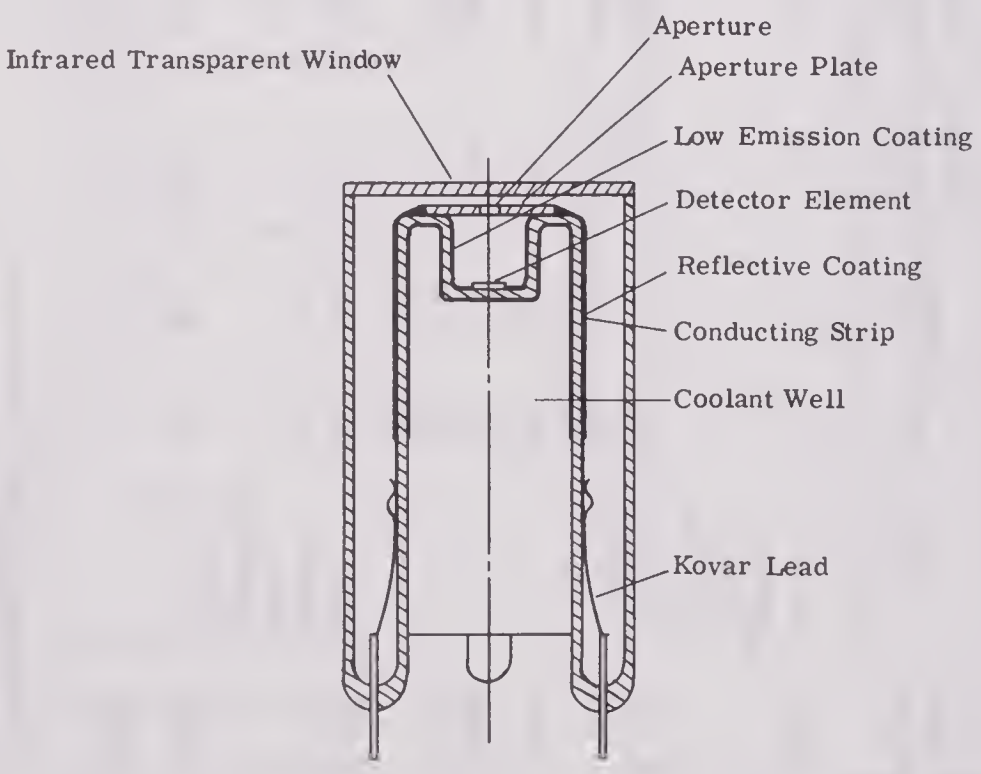


FIG. 12-29. Construction of a cooled aperture detector.

to ensure that the detector views a relatively cool background. The inside wall of the dewar and the face of the aperture plate is usually coated with a material of good reflection characteristics to reduce heat leak from radiation. Cooling the detector enclosure and the aperture plate decreases their background radiation, and also results in significant changes in several detector parameters, as indicated by the investigations on PbSe surfaces [15], and on PbS surfaces [16].

The contribution of optical system filters to the background radiation can be significantly decreased by lowering the temperature of the filter. This can be accomplished most conveniently by placing the filter inside the dewar between the window and the detector (see Chapter 11).

12.3. Space-Environment Cooling Systems [17]

Space cooling systems require extreme reliability, often coupled with a requirement for long periods of unattended operation. They are particularly restricted as to size and weight. The space environments of high vacuum and zero gravity create problems of liquid behavior. Development of a cooling system based on the sublimation of a solid coolant into the high vacuum of space shows considerable promise since the problems associated with vapor liquid separation while venting in a zero gravity field are avoided. This cooling system consists of a solidified gas or liquid, an insulated container, an evaporation path to space, and a conduction path from the coolant to the device being cooled (Fig. 12-30). The operating temperature obtainable with this system depends upon the choice of coolant, the pressure maintained in the system, and the heat load. By varying the vapor flow rate, which in turn regulates the back pressure and temperature of the effluent flow, a specific operating temperature can be maintained. The system's operating time depends on the amount of coolant and the heat load.

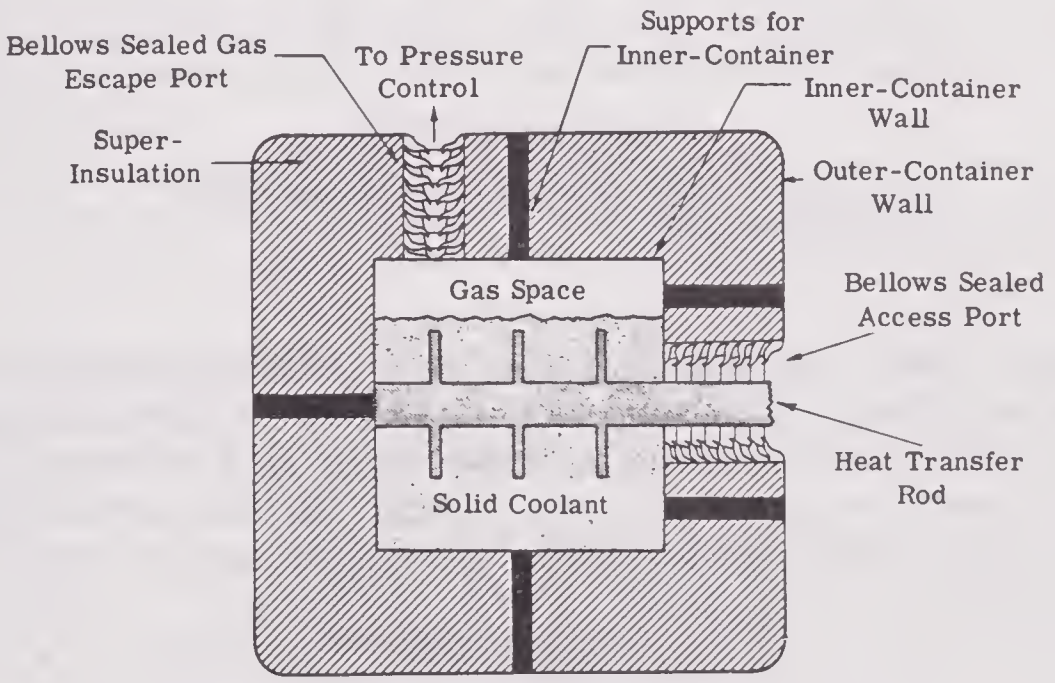


FIG. 12-30. Space environment coolant system [17].

12.3.1. Operating Principles. The space environment cooling system is based on the pressure dependency of the temperature of a solid in equilibrium with its vapor. The vapor pressure *versus* temperature plot for methane is shown in Fig. 12-31. For an operating temperature of 77°K using solid methane, the vapor pressure must be maintained at a value of 9.0 torrs. Addition of heat vaporizes the solid coolant and would cause the vapor pressure, and thus the temperature, to increase. The temperature and pressure are maintained constant by venting this vaporized coolant to the

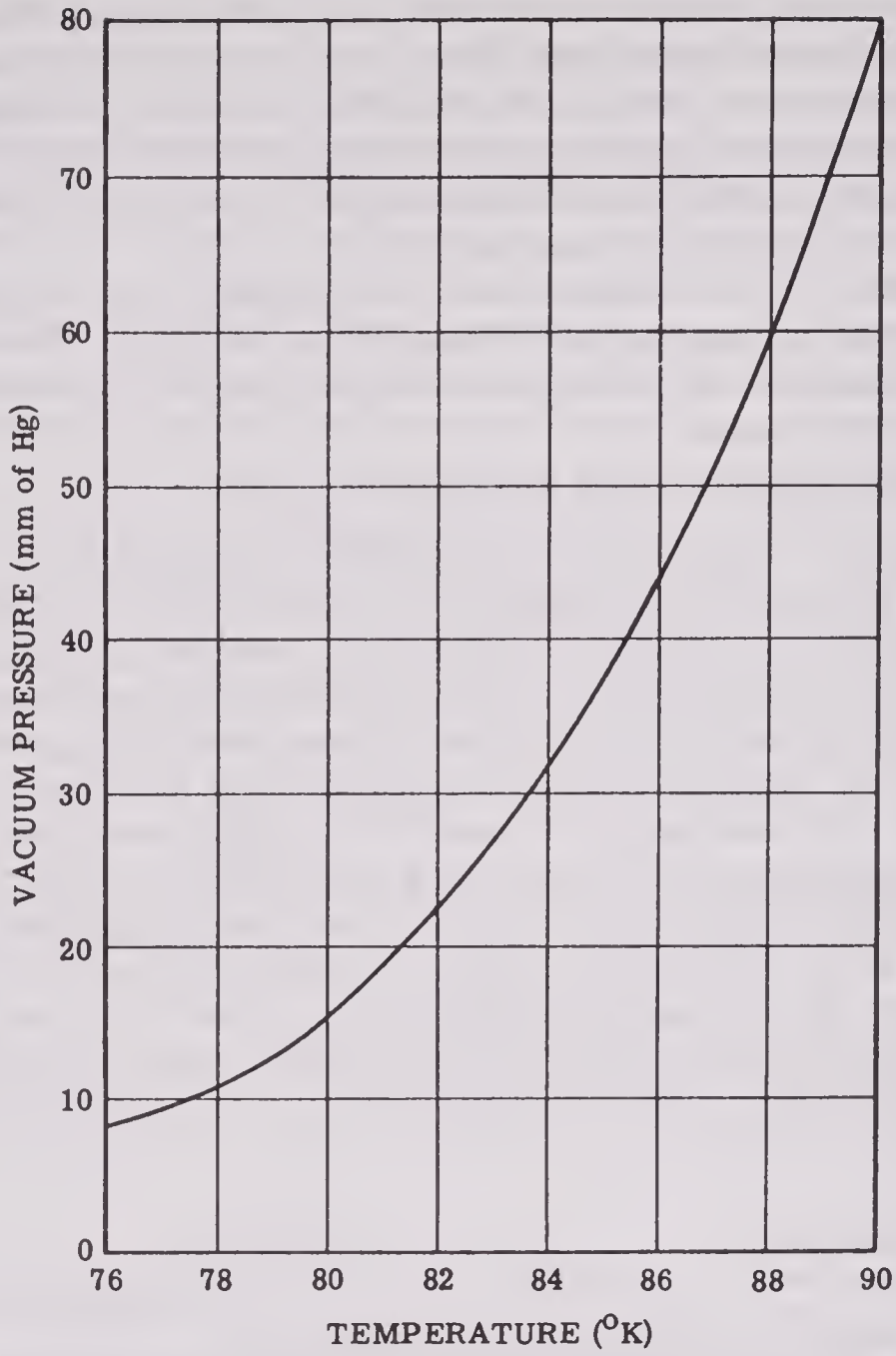


FIG. 12-31. Vapor pressure *versus* temperature of solid methane [17].

high vacuum of space. The quantity of heat that a given weight of solid coolant removes in vaporizing is equal to the sum of the heat of fusion and the heat of vaporization. For example, to sublimate 1 g of solid hydrogen, which has a heat of fusion of 0.0544 joule/kg and a heat of vaporization of 0.452 joule/kg, 506 joules of heat input are required. The rate at which the solid vaporizes per unit surface area is given by

$$Q = \frac{5.8\,P}{100}\sqrt{M/T} \tag{12-31}$$

where Q = rate in g sec^{-1} cm^{-2}

P = vapor pressure at temperature T in torrs

T = temperature in °K

M = molecular weight of the coolant

12.3.2. Design. The main factors entering into the design are the choice of coolant, surface-to-volume ratio of the container, and insulation requirements of the container and support structures.

The temperature and pressure ranges for several solid coolants are given in Table 12-9. Solid hydrogen is most promising because of the wide temperature range available. Systems have been conceived using solid hydrogen for cooling up to 140°K. At a higher temperature of 200°K, solid ammonia is another promising refrigerant because of its high density, which results in a smaller and lighter storage container.

TABLE 12-9. TEMPERATURE AND PRESSURE RANGES OF SOLID COOLANTS [17]

Coolant	*Temperature Range* (°K)	*Pressure Range* (torrs)*
Methane	90-67	80-1
Argon	83-55	400-1
Carbon monoxide	68-51	10-1
Nitrogen	62-47	73.6-1
Neon	24-16	260-1
Hydrogen	14-10	56-2

*Below this pressure only a 1° to 3° temperature drop takes place.

A cylinder with its length L equal to its diameter D is used as the coolant container since it has the lowest surface-to-volume ratio of any simple configuration. Superinsulation (Table 12-2) having a low thermal conductivity of 0.5 μw (cm °K)$^{-1}$ is used as the insulating material. Heat leak through the support structures is reduced by using materials of low thermal conductivity and small cross sectional area with the greatest possible length.

The main source of heat in the system is the heat leak into the coolant container. This is the critical element in obtaining optimum system weight as a function of operating time. In particular, the heat leak through the seals must be a minimum. These seals surround the heat transfer rod from the detector to the coolant and provide a path to the pressure control. Bellows are used as seals instead of straight tubes because of the greater heat-path length they provide for the same linear thickness. A 2.54×10^{-4} m-thick stainless-steel bellows with a heat-path length five times the insulation thickness is considered suitable.

This design provides a methane cooler with a one-year operating temperature of 77°K that weighs 13.8 kg, whereas a hydrogen cooler with a one-year operating temperature of 13°K weighs 31.7 kg. If a bellows seal of thinner material (5.08×10^{-5} m) and a heat-path length 20 times the insulation thickness is used, weight reductions of 9.06 kg for a methane cooler and 20.2 kg for a hydrogen cooler can be achieved. Thus, reducing the heat leak through the insulation or the ports provides a more efficient cooler in which the heat load of the detector is the controlling factor.

12.3.3. Practical Cooling Systems. The results of calculations of system weight using hydrogen, neon, nitrogen, carbon monoxide, argon, and methane are given in Table 12-10. Coolers using hydrogen and methane have the lowest weight for their respective temperature ranges. System weight and volume as a function of operating time for a solid-methane cooler are shown in Fig. 12-32 and 12-33, respectively. System weight and volume as a function of operating time for a solid-hydrogen

TABLE 12-10. COOLANT AND INSULATION WEIGHTS FOR CYLINDRICAL CONTAINER ($L=D$) AND ONE-YEAR OPERATION [17]

Coolant	*Coolant Temperature* (°K)	*Weight of Coolant and Insulation* (kg)
Hydrogen	12	29.8
Neon	24	53.9
Nitrogen	61	28
Carbon monoxide	68	26.6
Argon	84	30
Methane	88	11.9

Detector head load = 100 mw
Outer container temperature = 300° K

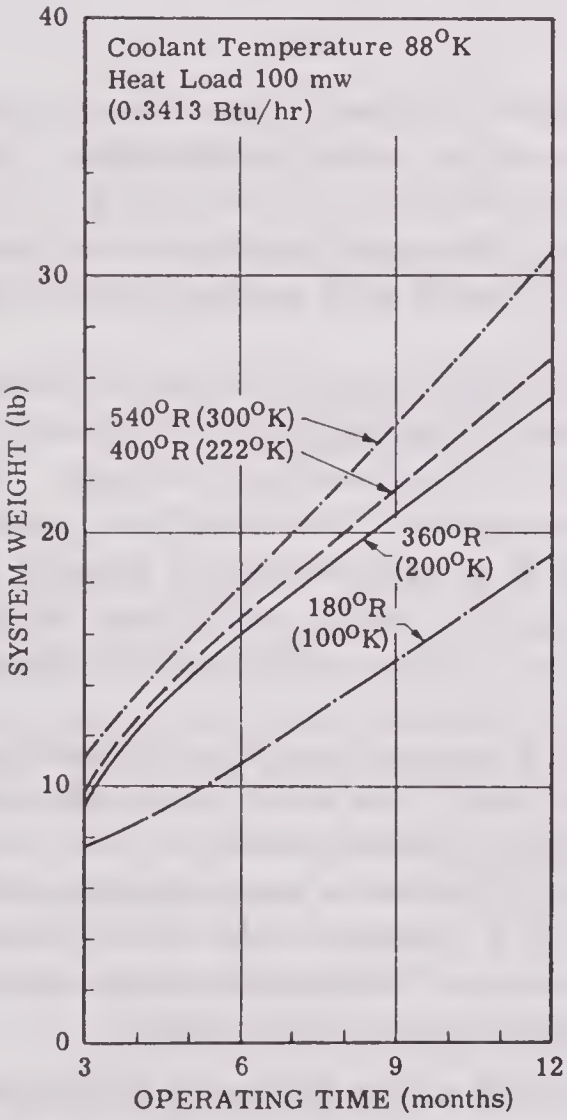


FIG. 12-32. System weight as a function of operating time at various outer container temperatures for methane coolant [17].

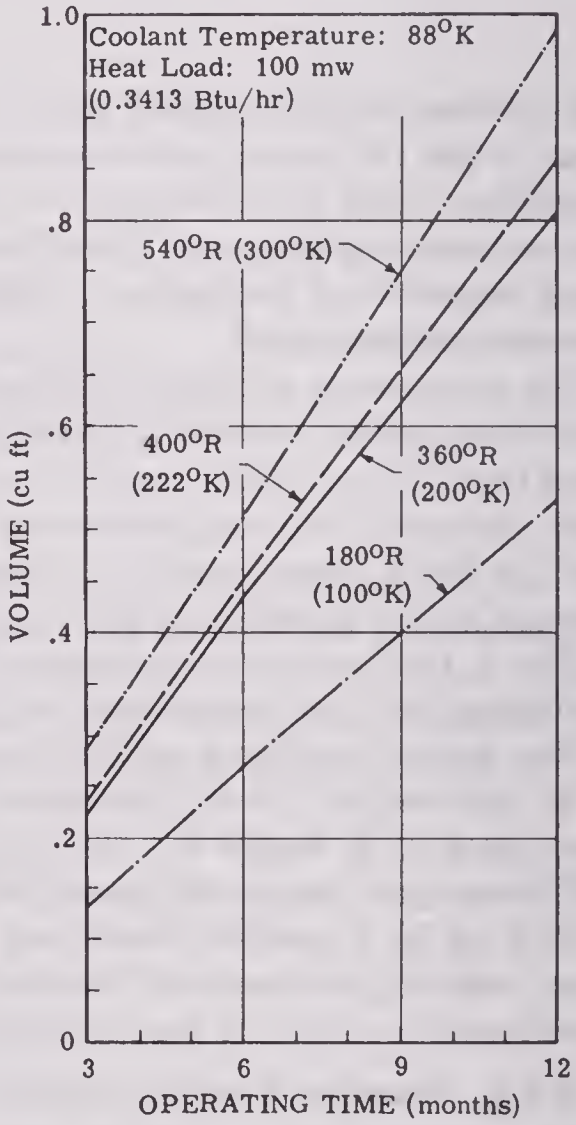


FIG. 12-33. System volume as a function of operating time at various outer container temperatures for methane coolant [17].

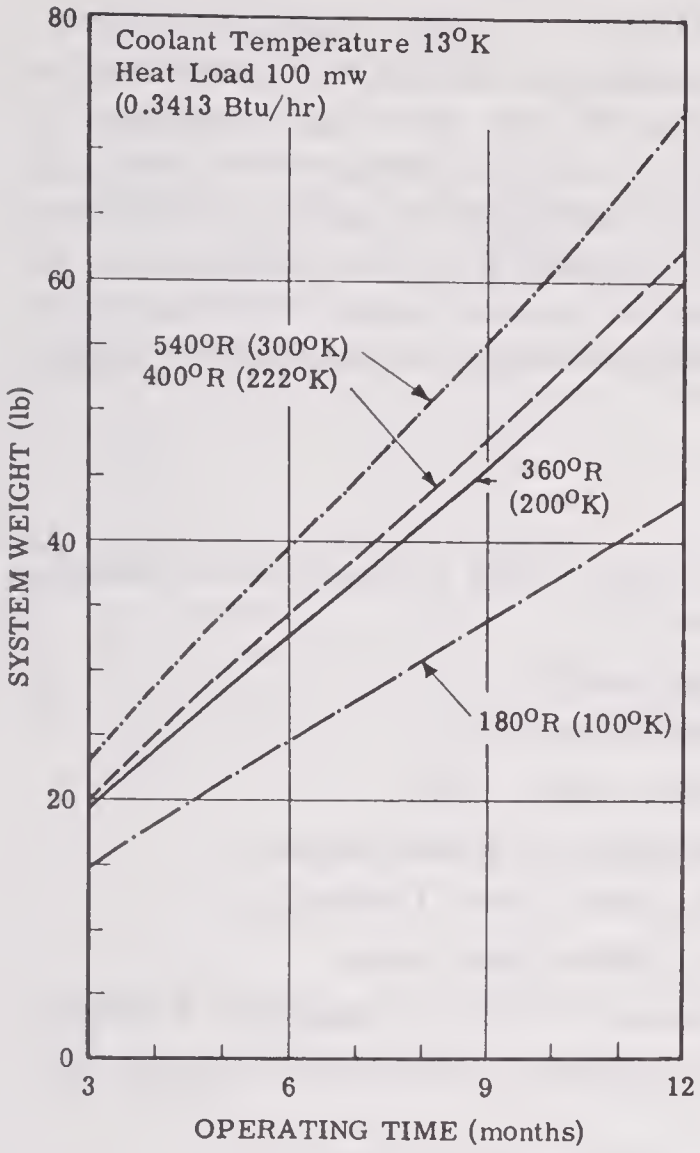


FIG. 12-34. System weight as a function of operating time at various outer container temperatures for hydrogen coolant [17].

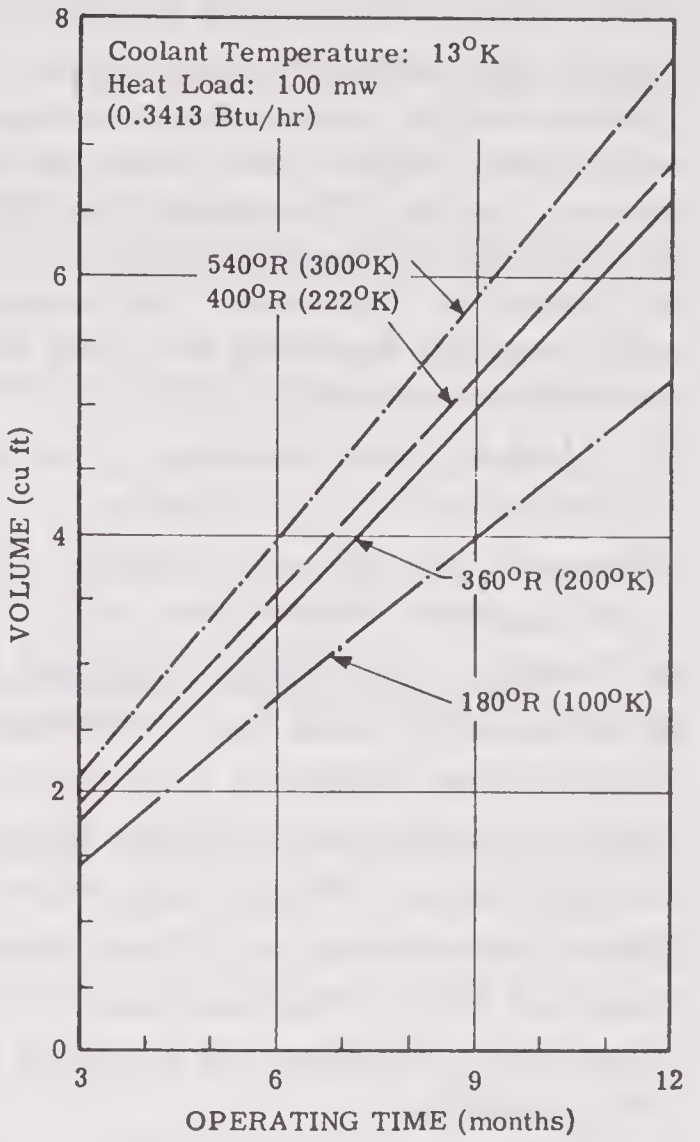


FIG. 12-35. System volume as a function of operating time at various outer container temperatures for hydrogen coolant [17].

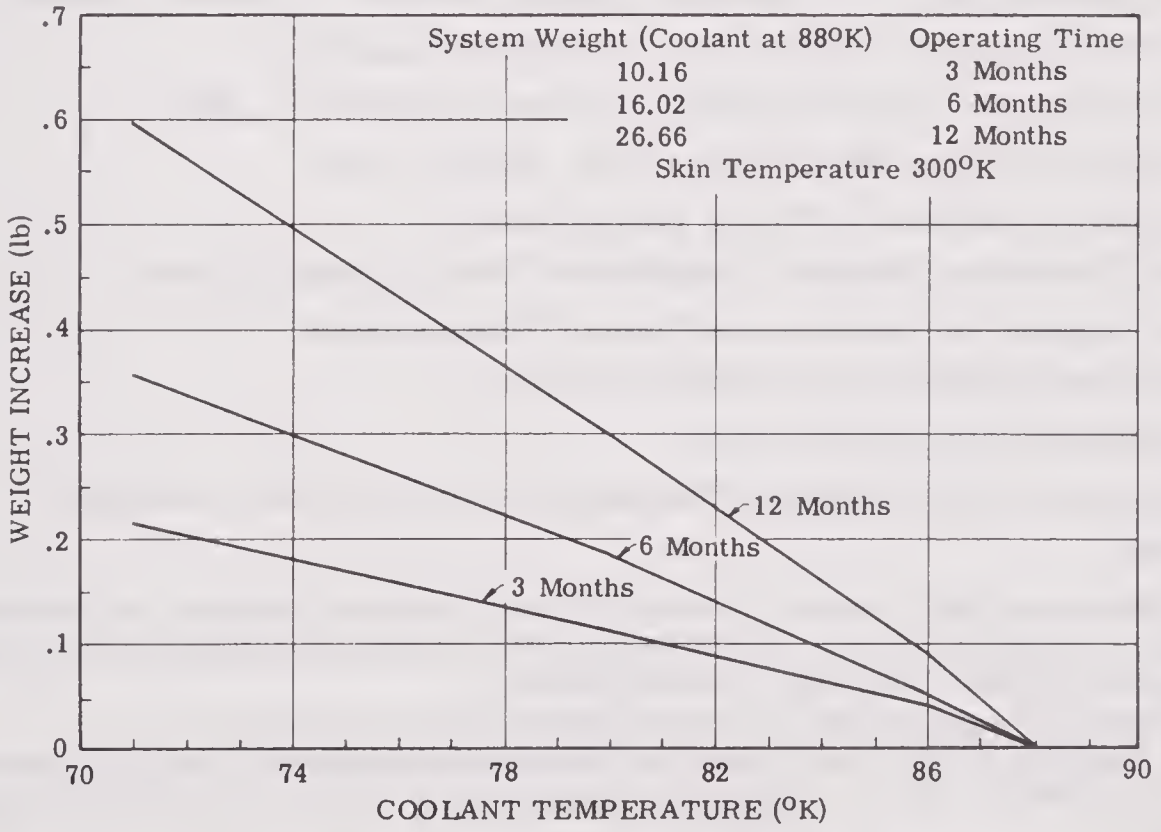


FIG. 12-36. Increases of methane cooling system weight with decrease in coolant temperature [17].

cooler are shown in Fig. 12-34 and 12-35, respectively. The effects of ambient temperature are also indicated on these plots. The weight and volume of the methane and hydrogen cooling systems were calculated using only one operating temperature for each coolant. Figure 12-36 shows the results of calculations made to determine the change in weight of the system when different temperatures are used. The weight of the container, the pressure-control, and the heat-transfer rod were neglected in the calculations for Fig. 12-36. The container weight does not exceed 0.453 kg for any configuration or operating time, and the pressure-control and transfer-rod weights were estimated to total 1.812 kg.

12.4. List of Manufacturers

Aerojet-General, Azusa, California

AiResearch Manufacturing Division, Garrett Corp., 9851 S. Sepulveda Boulevard, Los Angeles 45, California

Air Products and Chemicals, Allentown, Pennsylvania

Arthur D. Little, Acorn Park, Cambridge, Massachusetts

Cryogenerators, Division N.A. Philips Co., Ashton, Rhode Island

Energy Conversion Inc., 336 Main Street, Cambridge 42, Massachusetts

Fairchild Stratos, 1800 Rosecrans Avenue, Manhattan Beach, California

Hofman Laboratories, Inc., 5 Evans Terminal, Hillside, New Jersey

Honeywell, Military Products Group, 1915 Armacost Avenue, Los Angeles 25, California

International Telephone and Telegraph Corp., Industrial Laboratories Division, Fort Wayne, Indiana

Jepson Thermoelectric, Inc., 139 Nevada Street, El Segundo, California

Linde Co., Div. Union Carbide Corp., 270 Park Avenue, New York 17, New York

Malaker Laboratories, Inc., West Main Street, High Bridge, New Jersey

Martin Company, Electronic Systems and Products Division, Baltimore 3, Maryland

Materials Electronics Products Corp., 990 Spruce Street, Trenton, New Jersey

Needco of America, Inc., 5770 Andover Avenue, Montreal 9, Quebec, Canada

Norden, Division United Aircraft Corp., Norwalk, Connecticut

Ohio Semiconductors, 1205 Chesapeake Avenue, Columbus 12, Ohio

Pesco Products Division, Borg-Warner Corp., Bedford, Ohio

Philco, Lansdale Division, Lansdale, Pennsylvania

Radiation Electronics, Division, Comptometer Corp., Chicago, Illinois

Raytheon, Spencer Laboratories, Burlington, Massachusetts

Santa Barbara Research Center, Goleta, California

Texas Instrument Inc., Dallas, Texas

Westinghouse, Air Arm Division, Friendship Airport, Baltimore, Maryland

References

1. P. R. Barker and W. L. Brown, "Cooling Devices for Infrared Detectors," in *Infrared Quantum Detectors,* Institute of Science and Technology, The University of Michigan, Ann Arbor, Mich., Rept. No. 2389-50-T (July 1961) (CONFIDENTIAL).
2. Data Sheet, "Cryogenic Data," Cryogenic Engineering Co., Denver, Colo.
3. J. R. Kettler and P. L. Rice, *Closed Cycle Helium Refrigeration for Cooling Electronic Components,* AE-1991-R, Rev. 1. AiResearch Manufacturing Division, Garrett Corp., Los Angeles (1962).
4. *Instruction Manual for Nitrogen Pressure Generator,* Barnes Engineering Co., Stamford, Conn. (1961).

5. Sales Brochure, *Infrared Cooling*, SER-24-1, AiResearch Manufacturing Division, Garrett Corp., Los Angeles (1962).
6. Sales Brochure, *Infrared Components*, Santa Barbara Research Center, Goleta, California (1962).
7. Sales Brochure, *Moisture Trap for Drying High Pressure Nitrogen Gas*, ITT Components and Instrumentation Lab., Fort Wayne, Ind. (1961).
8. R. H. Vought, "Peltier Cooling," in *Infrared Quantum Detectors*, Institute of Science and Technology, The University of Michigan, Ann Arbor, Mich., Rept. No. 2389-50-T (July 1961) (CONFIDENTIAL)
9. B. L. Worsnop, *Applications of Thermoelectricity*, Wiley, New York, 8-106 (1960).
10. R. R. Heikes and R. W. Ure, Jr., *Thermoelectricity, Science and Engineering*, Interscience, New York, **15**, 458-517 (1961).
11. Sales Brochure, *Thermoelectric Cascade Cooler 094438-010*, Pesco Products Division, Borg-Warner Corp., Bedford, Ohio (1962).
12. M. B. Grier, *Proc. IRE*, **47** (1959).
13. F. E. Jaumet, *Proc. IRE*, **46**, 538 (1958).
14. H. J. Goldsmid, *Applications of Thermoelectricity*, Wiley, New York, N. Y. (1960).
15. R. M. Talley, T. H. Johnson, and D. E. Bode, *Proc. IRIS*, **4**, No. 4 (1959).
16. J. J. McArdle, *Proc. IRIS*, **6**, No. 1, 107, Infrared Industries, Waltham, Mass. (1961).
17. A. I. Weinstein, A. S. Friedman, and U. E. Gross, *Proc. IRIS*, **7**, No. 2, 187-191, Aerojet-General Corp., Azusa, Calif. (1962).

Chapter 13

FILM

Allan L. Sorem

Eastman Kodak Company

Gwynn H. Suits

The University of Michigan

CONTENTS

13.1. Introduction 570
13.2. Available Infrared Films and Plates 570
13.3. Hypersensitizing 570
13.4. Definition of Density and Exposure 573
13.5. Sensitometric Characteristics 573
13.6. Spectral Sensitivity and Filter Transmittances 574
13.7. Reciprocity Characteristics 574
13.8. References for Additional Details 574

13. Film

13.1. Introduction

Most of the characteristics of infrared photographic film and plates are not unlike those of visible light films and plates. The primary distinction is that the infrared emulsion responds to radiation in the near-infrared spectral region as well as to some part of the visible and ultraviolet.

Infrared films can be handled like similar conventional materials, except that extra precautions should be observed to make sure that the films will not be fogged by unsuspected stray infrared radiation. Loading and unloading of cameras, magazines, or cassettes should preferably be done in total darkness. Certain woods, hard rubbers, plastics, and black papers are not opaque to near infrared radiation. Thus, accidental fogging of infrared film can occur when improper protective covering is used for the plates. Metal foil on black paper is a dependable covering. If a safe-light is needed during processing, it must be used with a filter specifically recommended for infrared materials, since most safe-lights transmit infrared freely. The Kodak Safe-Light Filter, Wratten Series 7, is recommended.

When infrared film is used in equipment designed primarily for visible-light photography, some additional precautions should be noted. The lens focal positions indicated by barrel markings or range finders on cameras designed for visible-light photography do not apply accurately when infrared film is used. The longer wavelength rays are refracted less by lenses so that the effective focal length of the lens is about 0.5% longer than the visible-light focal length. In addition, aberration corrections which are made to optimize the sharpness of the image in the visible portion of the spectrum are not generally optimum for the infrared portion. Exposure meters designed for visible light use do not respond to near-infrared radiation. Hence, these meter readings are not useful as exposure indicators unless the ratio of visible to infrared radiation is known beforehand.

13.2. Available Infrared Films and Plates

Infrared films are available in roll form for still and motion-picture cameras, and in sheet form. In addition, spectroscopic plates and films with spectral sensitivity in the near-infrared region are available. Table 13-1 lists brief descriptions of these materials as an indication of the types of films currently available and of the forms in which they are sold through dealers.

13.3. Hypersensitizing

The speeds of some infrared films and plates can be increased by hypersensitizing them. For best results this should be done just before exposure, and in any case the films and plates should be used within a few days after hypersensitizing. This procedure is essential for Kodak spectroscopic films and plates having M, N, and Z sensitizings.

Films or plates to be hypersensitized should be bathed in a weak solution of ammonia made by diluting 4 parts of 28% ammonia (the strongest available commercially) with

TABLE 13-1. INFRARED FILMS AND PLATES

Films and Plates	*Characteristics and Use*	*Available Sizes*	*Source*	*Remarks*
Kodak Infrared Film (Sheet)	Moderately high-contrast, antihalation film for distant haze penetration	All regularly listed sheet-film sizes	Eastman Kodak dealers	With variations in exposure and processing, can be used in certain types of scientific, medical, and documentary photography, photomicrography, and photomechanical work
Kodak Infrared Film (35 mm)	Moderately high-contrast film for distant landscape photography. Records detail ordinarily obscured by atmospheric haze.	(1) 20-exposure magazines (2) 50-ft rolls	Eastman Kodak dealers	Also useful in special applications requiring infrared records
Eastman Infrared Neg. Film, Type 5210	Professional 35-mm motion picture negative film	100-,200-,400-,1000-, and 2000-ft rolls	Motion Picture Film Dept., Eastman Kodak Co., Rochester 4, N.Y.	
Kodak High-Speed Infrared Film	16-mm and 35-mm motion picture film	100-foot rolls	Eastman Kodak dealers	
Kodak Infrared Aerographic Film		All regularly listed aerial-film sizes		
Kodak Ektachrome Infrared Aero Film	Three-layer reversal color material, orginally for camouflage detection. All layers are blue-sensitive. One layer, which is sensitive to green light, produces a yellow image, another layer is sensitive to red light and produces a magenta image, and the third layer, sensitive to infrared to about 0.9 μ, produces a cyan image.	Aerial film sizes	Eastman Kodak dealers	When processed in the Kodak E-3 reversal color process, a false-color image is produced. Healthy green foliage and other high-infrared-reflecting objects are represented by shades of red and pink. Most green paints and diseased foliage represented as bluish.

TABLE 13-1. INFRARED FILMS AND PLATES (*Continued*)

Films and Plates	*Characteristics and Use*	*Available Sizes*	*Source*	*Remarks*
Kodak Spectroscopic Films and Plates		Films – 35-mm and 16-mm 100-ft rolls Plates 1×3″ 8×10″ 2×10″ 6.5×9 cm 3-1/4×4-1/4″ 9×12 cm 5×7″ 16×16 cm	Special Sensitized Products Sales Div., Eastman Kodak Co., Rochester 4, N.Y. (special order basis)	All except Type 103a-U must be hypersensitized (see Sec. 13.4).
Type I-M	Medium-contrast medium resolving power plate or film. Coarse grain. Sensitive to about 1 μ.			
Type I-N	Similar to Type I-M, but sensitive to about 0.9 μ, with a peak sensitivity at 0.8 μ.			
Type IV-N	Extremely high contrast, very high resolving power, fine grain. Same spectral range as Type I-N, but much lower sensitivity.			
Type I-Z	Medium-contrast, medium resolving power plate or film. Coarse grain. Sensitive to about 1.2 μ, with a peak sensitivity at 1.1 μ.			
Type 103a-U	Medium contrast, moderately low resolving power plate or film. Moderately coarse grain. Sensitive to about 780 mμ, with a peak sensitivity at 740 mμ.			

100 parts of water. *Concentrated ammonia is a dangerous irritant and should be handled cautiously.* The diluted ammonia solution should preferably be used at a ture of about 40°F but never at a temperature exceeding 55°F. The films or plates should be bathed for about 3 minutes. The material should then be bathed for 2 to 3 minutes in methyl or ethyl alcohol and dried as rapidly as possible in a stream of cool, dust-free air from a blower or fan.

Kodak Spectroscopic Plate and Film Type I-N can be hypersensitized by using plain water instead of the ammonia solution described above.

Care must be exercised in hypersensitizing to prevent streaking. In drying, particular care must be taken to keep dust from the emulsion surface, to prevent spotting.

13.4. Definition of Density and Exposure

The exposure and development of photographic film produces an image consisting of areas having different transmittances, depending on the number and size of the silver grains present. If the transmittance is measured by the ratio of the intensity of the undeviated light passing through the plate or film to that of the incident collimated light on the back, then the transmittance is called the *specular transmittance.* If the transmittance is measured by the ratio of the intensity of the undeviated and scattered light together to that of the incident collimated light, then the transmittance is called the *diffuse transmittance.* The opacity, O, is defined as the reciprocal of the transmittance. The density, D, is defined as $\log_{10} O$. Thus, depending upon the manner of measurement of transmittance, there is a corresponding specular and diffuse density. The density values used in the data of this chapter are diffuse density values.

The exposure, E, may be expressed either as the time integral of the illuminance, I, or the time integral of irradiance, H, so that

$$E \text{ (photometric)} = C \int_0^t \int_{\lambda_1}^{\lambda_2} V(\lambda) H(\lambda)\, d\lambda\, dt$$

and

$$E \text{ (radiometric)} = \int_0^t \int_{\lambda_1}^{\lambda_2} H(\lambda)\, d\lambda\, dt$$

where $V(\lambda)$ is the relative visibility curve and C is a conversion factor. In photographic literature, when E is expressed as a photometric quantity, the units of E are customarily meter-candle-seconds. When E is expressed as a radiometric quantity, the units of E are customarily ergs per square centimeter.

13.5. Sensitometric Characteristics

The curves presenting density as a function of $\log_{10} E$ are known as the H and D curves (after Hurter and Driffield), or usually just the characteristic curves (see Fig. 13-1 to 13-8). The standard source of radiation for obtaining these curves is an artificial source which is made to provide an irradiance the spectral distribution of which is close to that provided by a 6100°K blackbody in the restricted spectral interval of interest. This source is intended to simulate average daylight (sunlight plus skylight). From the illuminance from this source one may derive a unique irradiance in the infrared spectral region. Although more direct use of E (radiometric) for characteristic curves of infrared film makes good sense, E (photometric) has been used in the literature and is a valid indicator of the magnitude of E (radiometric). When a filter is specified along with the characteristic curves, the effects of the filter losses are implicit in the characteristic curves. The filter is considered to be an integral part of the film.

It is important to point out that the characteristic curves are average properties and have a meaning in this regard similar to published characteristics of vacuum tubes. In addition, these characteristics hold only for the specified conditions of exposure and processing. Changes in exposing light quality or in processing will yield a different set of characteristic curves. The response of photographic materials to parameter variation tends to be nonlinear with almost every parameter. It cannot be too strongly stressed that, when photometric or radiometric measurements are to be made, the apparatus to be used must be calibrated for photographic response. Extreme care must be exercised to obtain reproducible processing conditions. When possible, calibrating exposures should be made adjacent to the areas to be measured. A wealth of information concerning the use and problems of photographic materials for photometry may be found in the literature of astronomy.

13.6. Spectral Sensitivity and Filter Transmittances

The sensitivity of a photographic film to monochromatic light is defined as the reciprocal of the monochromatic exposure (in erg cm^{-2}) required to produce a stated density above the fog level. The common logarithm of the sensitivity as a function of wavelength is given in Fig. 13-9 through 13-14. Notice that in Fig. 13-12 the sensitivity curve for $D = 0.3$ is not exactly a vertical translation of the curve for $D = 1.0$. This indicates that gamma is a function of wavelength as well as a function of development time. Transmission curves for some useful filters are found in Fig. 13-15.

13.7. Reciprocity Characteristics

The reciprocity law states that, when all other parameters are held fixed, the final density on the film depends only upon the produce of the illuminance due to the standard source and the exposure time. This law does not hold for extremes in exposure conditions. Very high intensity light of short duration and very low intensity light of long duration are both less effective in achieving a given density than moderate values of light and exposure times. The reciprocity characteristics are shown in Fig. 13-16 through 13-18, which are plots of log exposure in meter-candle-seconds *vs.* log illuminance (photometric irradiance), with constant exposure time lines indicated. The degree of deviation from a straight horizontal line shows the degree of reciprocity-law failure.

13.8. References for Additional Details

An important reference for more detailed discussion of infrared photographic materials, sources of illumination, and applications, is Walter Clark, *Photography by Infrared,* (Wiley, New York, 1946). A good discussion of photographic materials and materials and photometric concepts and terminology may be found in Hardy and Perrin, *The Principles of Optics,* (McGraw-Hill, New York, 1932). A chapter devoted to "Astronomical Photographic Photometry" may be found in F. E. Ross, *The Physics of the Developed Photographic Image* (Van Nostrand, New York, 1924).

Additional information about infrared photographic materials may be found in two Kodak publications: *Infrared and Ultraviolet Photography,* No. M-3, and *Kodak Plates and Films for Science and Industry,* No. P-9.

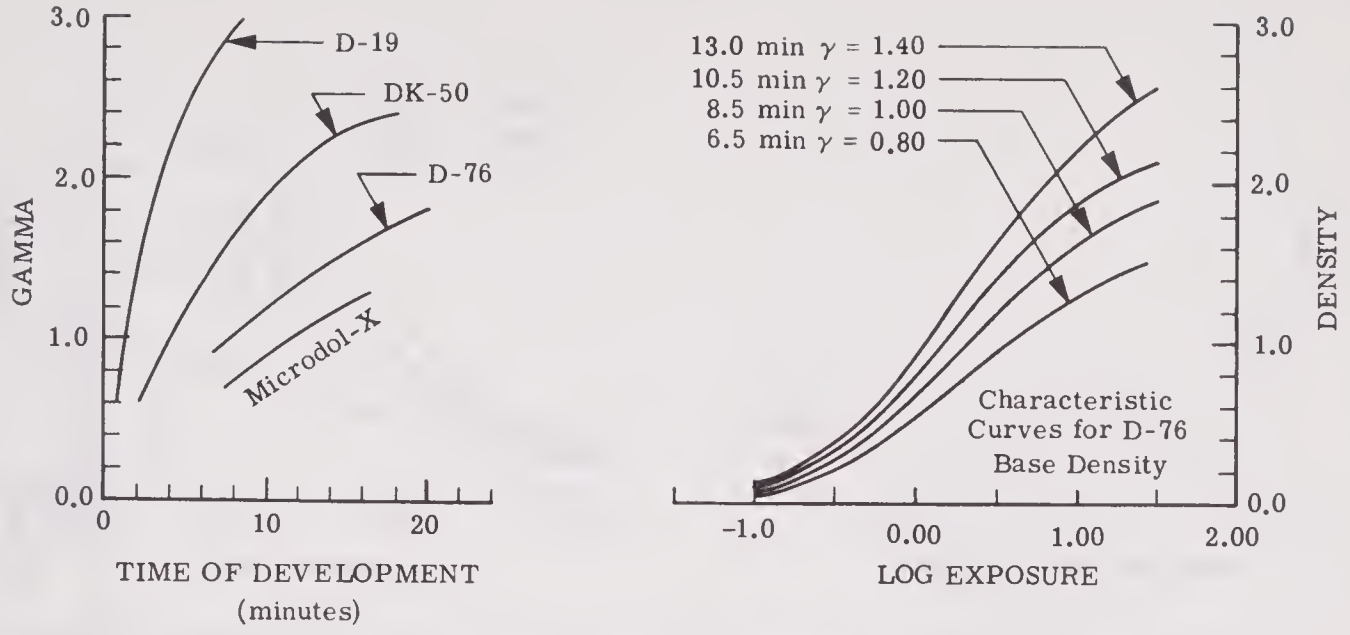


FIG. 13.1. Characteristics of Kodak Infrared Sheet Film exposed to sunlight through Wratten A filter No. 25, developed with intermittent agitation at 68°F (20°C).

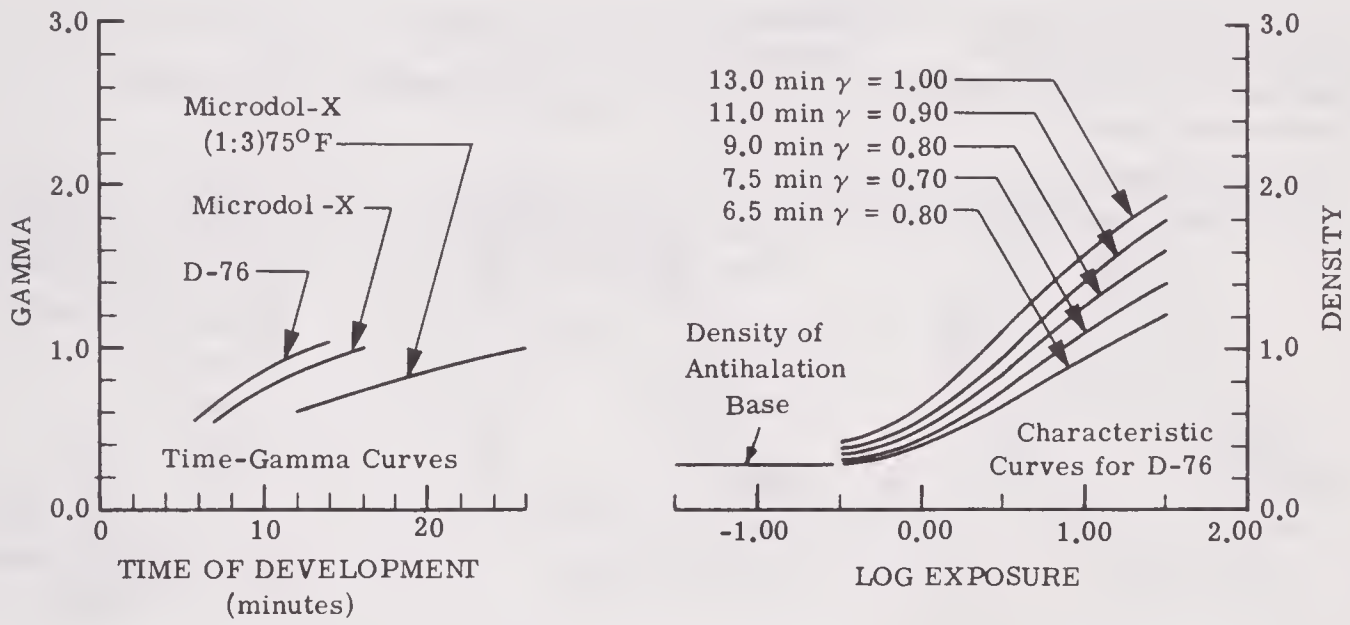


FIG. 13-2. Characteristics of Kodak Infrared Sheet Film (35 mm), exposed to sunlight through Wratten A filter No. 25, developed with intermittent agitation at 68°F (20°C).

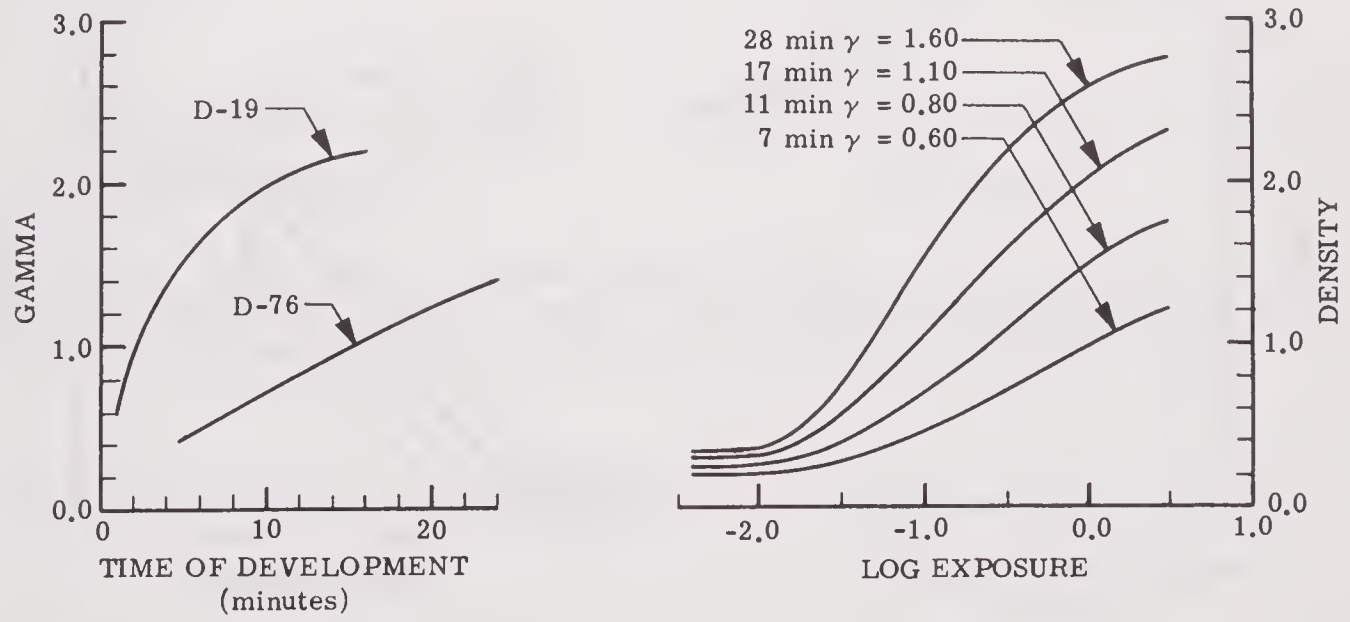


FIG. 13-3. Characteristics of Kodak High-Speed Infrared Film (H5218), exposed to sunlight through Kodak Wratten filter No. 25, developed in Kodak Developer D-76 at 68°F (20°C) with intermittent agitation.

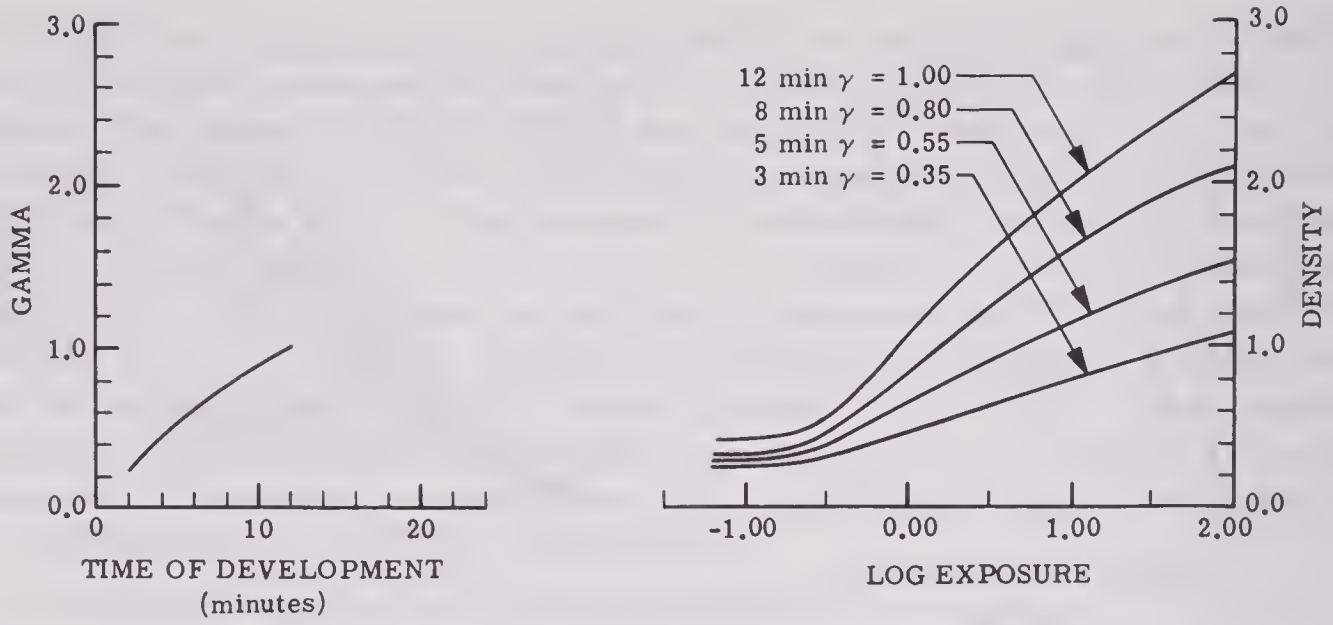


FIG. 13-4. Characteristics of Eastman Infrared Negative Film, Type 5210, exposed to daylight through Kodak Wratten filter No. 25, developed in Kodak Developer No. D-76 in a sensitometric machine.

Illuminant: Infrared (Daylight through Wratten #25 Filter)
Development: DK-50 at 68°F (20°C) in Sensitometric Machine

Illuminant: Infrared (Daylight through Wratten #25 Filter)
Development: D-19 at 68°F (20°C) in Sensitometric Machine

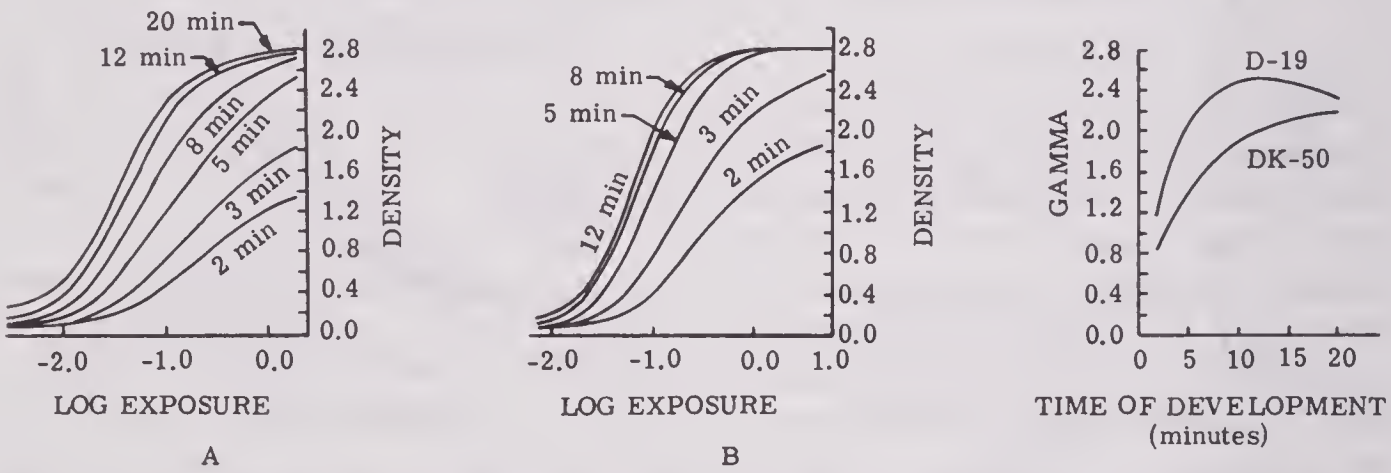


FIG. 13-5. Characteristics of Kodak Infrared Aerographic Film. Daylight through a No. 25 filter is the illuminant; for *A* the developer was DK-50 at 68°F (20°C), for *B* it was D-19 at 68°F (20°C)—both in a sensitometric machine.

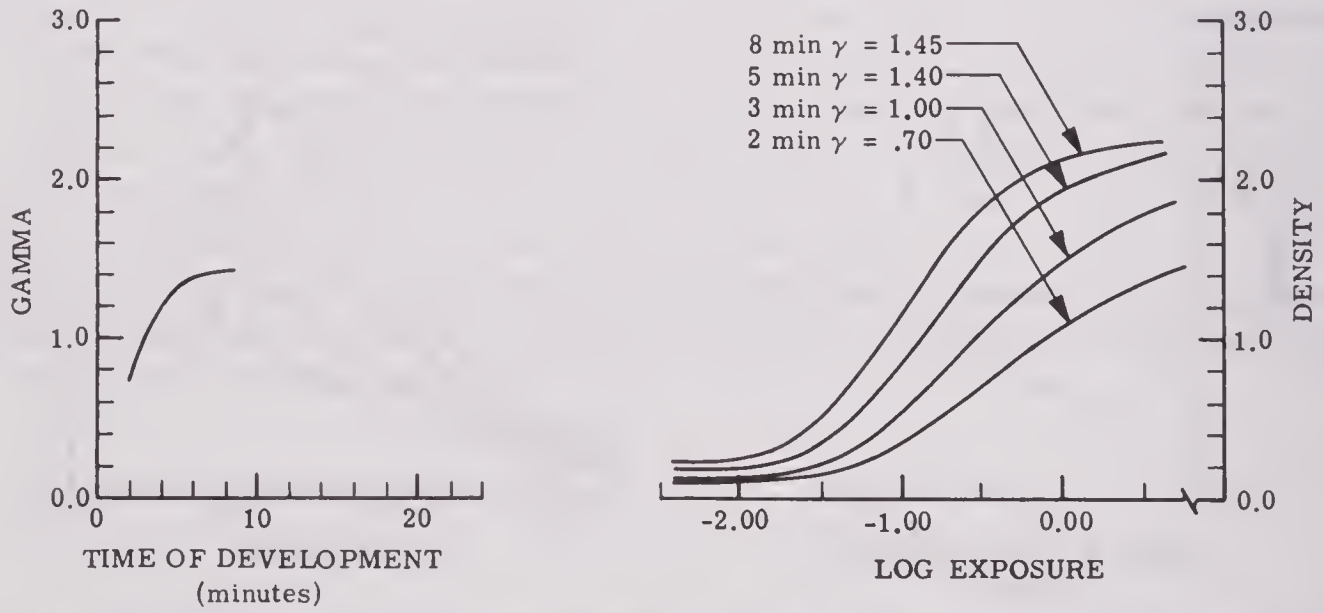


FIG. 13-6. Characteristics of Kodak Spectroscopic Plate and Film, Type I-N hypersensitized with water, exposed to daylight through Kodak Wratten filter No. 25, developed in Kodak Developer No. D-19 at 68°F (20°F) in a sensitometric machine.

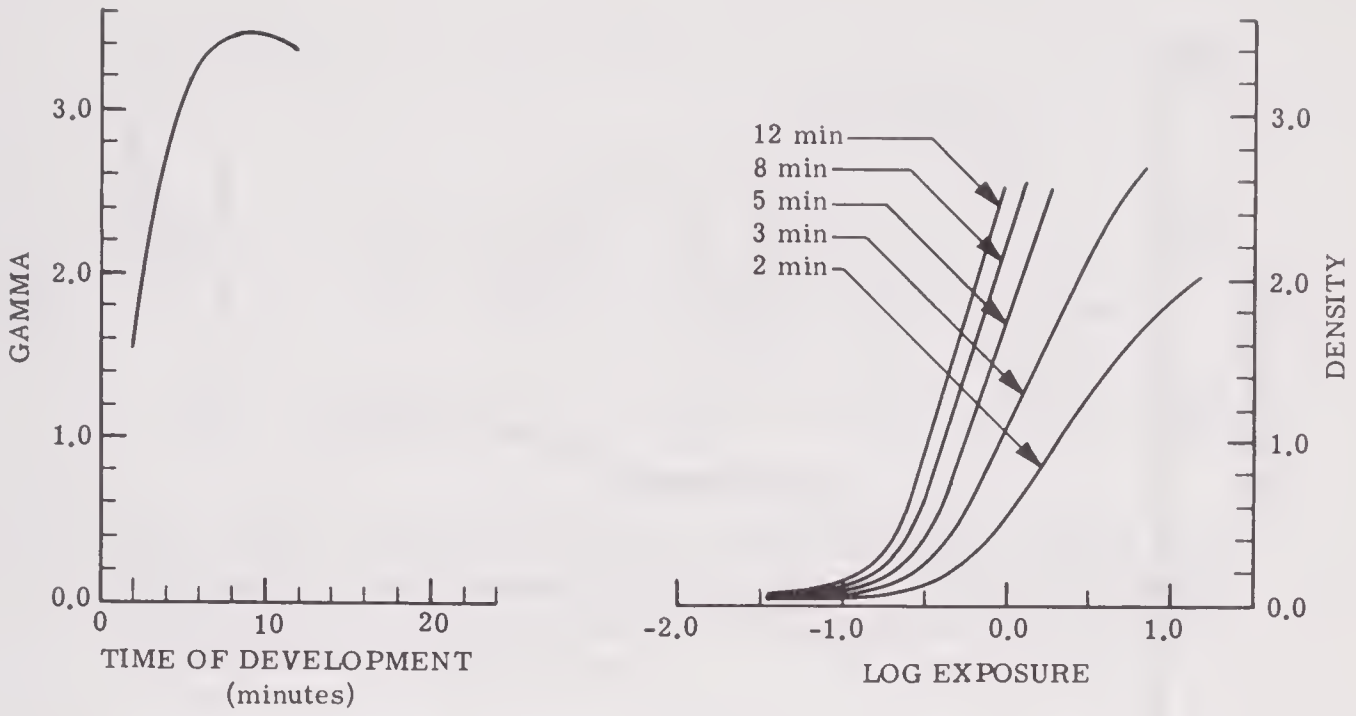


FIG. 13-7. Characteristics of Kodak Spectroscopic Plate and Film, Type IV-N, hypersensitized with ammonia, exposed to daylight through Kodak Wratten filter No. 25, developed in Kodak Developer No. D-19 in a sensitometric machine.

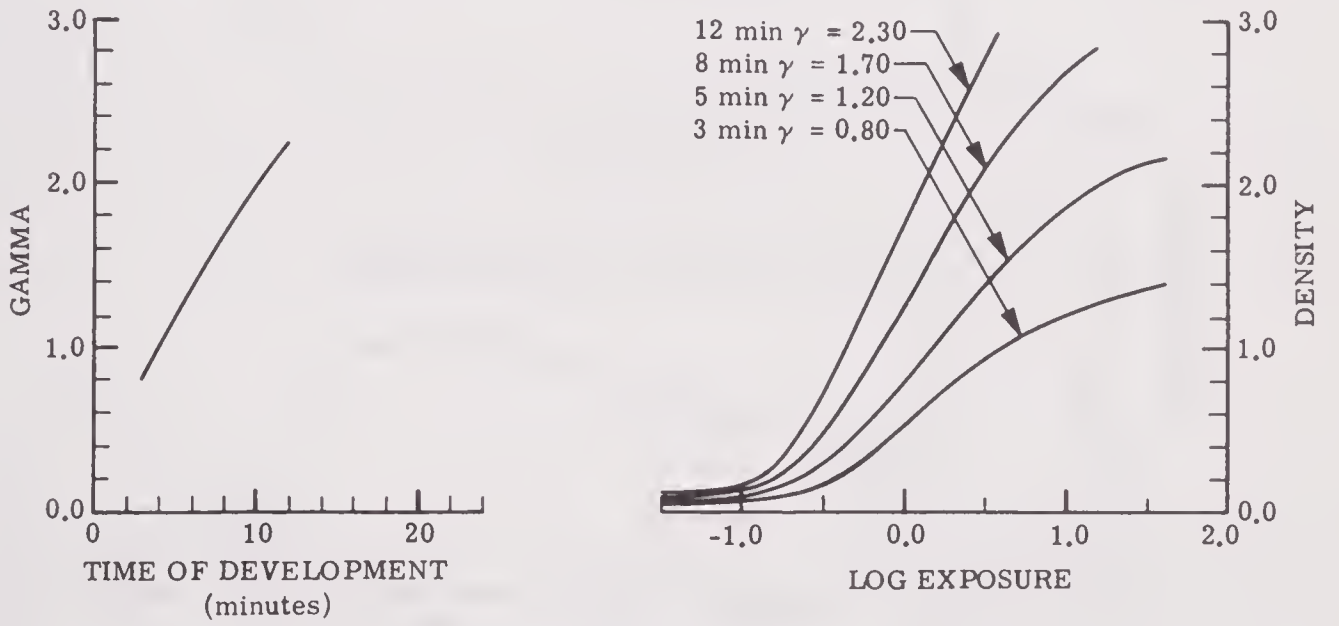


FIG. 13-8. Characteristics of Kodak Spectroscopic Plate and Film, Type IV-N, hypersensitized with ammonia, exposed to daylight through a Kodak Wratten filter No. 25, developed in Kodak Developer No. D-76 at 68°F (20°C) a sensitometric machine.

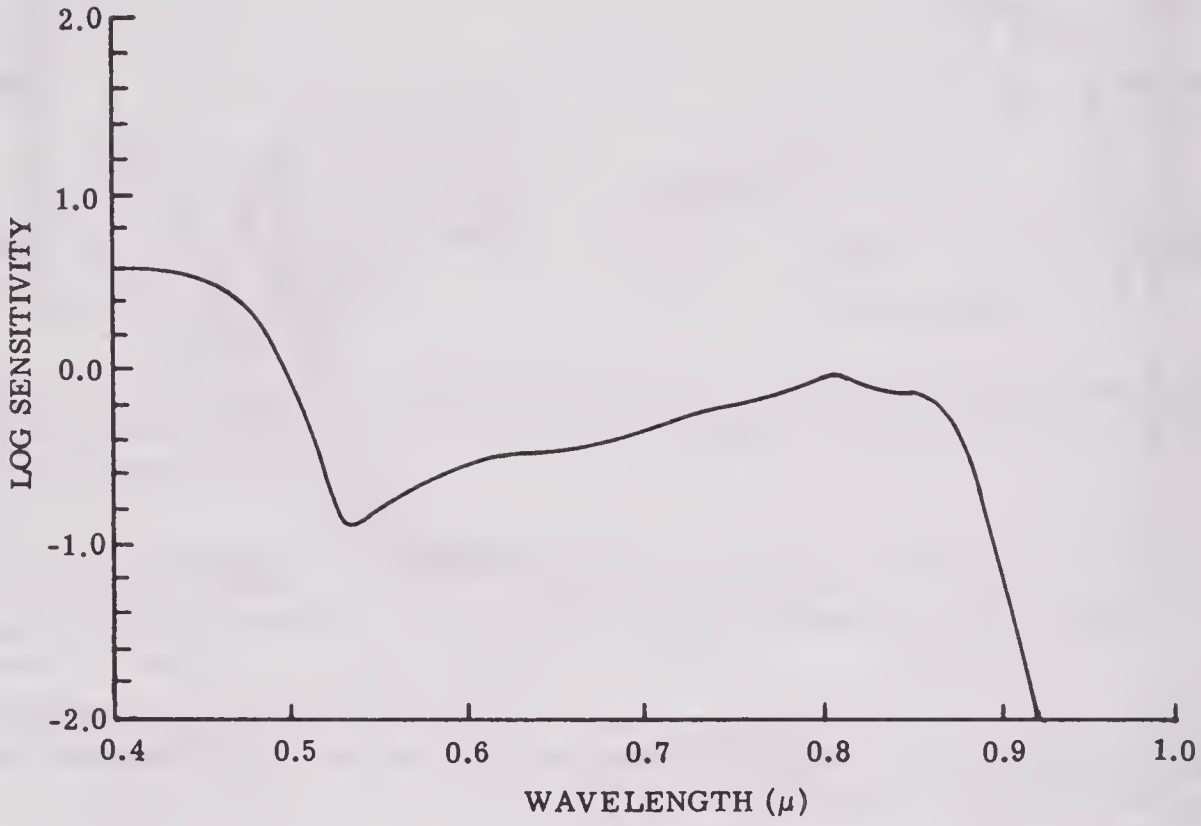


FIG. 13-9. Spectral sensitivity of Kodak Infrared Film developed in Kodak Developer DK-50 5 min at 68°F (20°F). $D = 0.6$ above gross fog.

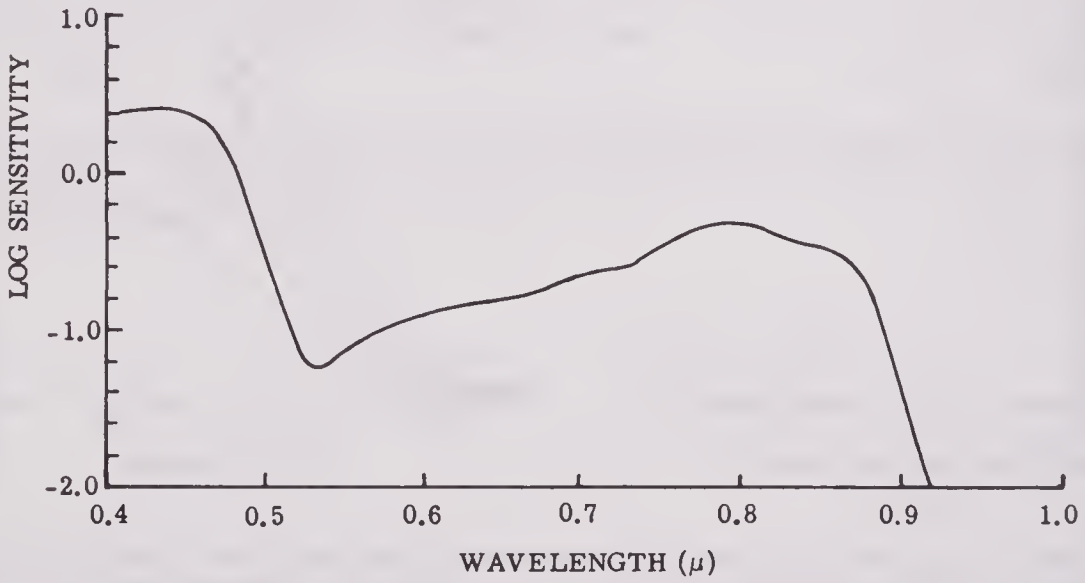


FIG. 13-10. Spectral sensitivity of Eastman Infrared Negative Film, Type 5210 developed in Kodak Developer D-76 9 min at 68°F (20°F). $D = 0.6$ above gross fog.

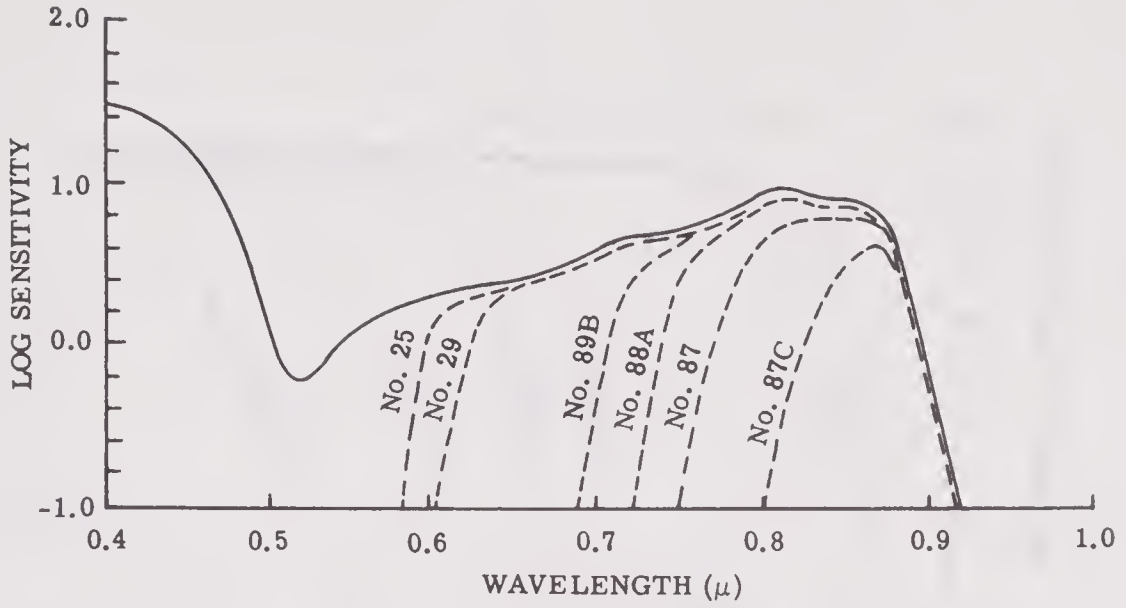


FIG. 13-11. Spectral sensitivity of Kodak High-Speed Infrared Film developed in Kodak Developer D-76 7-1/2 min at 67°F (20°F). $D = 0.6$ above gross fog. Dashed lines show effective sensitivities when the Kodak Wratten filters indicated are used.

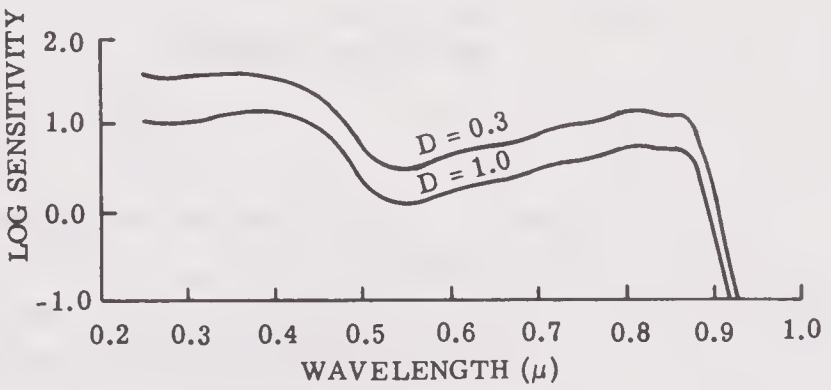


FIG. 13-12. Spectral sensitivity of Kodak Spectroscopic Plate and Film, Type I-N.

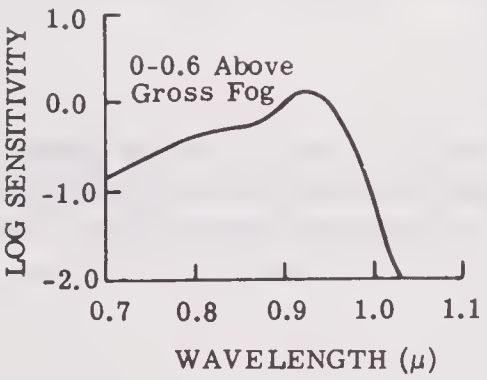


FIG. 13-13. Spectral sensitivity of Kodak Spectroscopic Plate and Film, Type I-M.

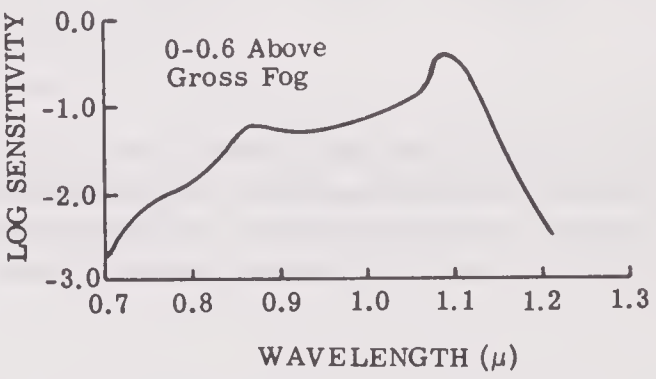


FIG. 13-14. Spectral sensitivity of Kodak Spectroscopic Plate and Film, Type I-Z.

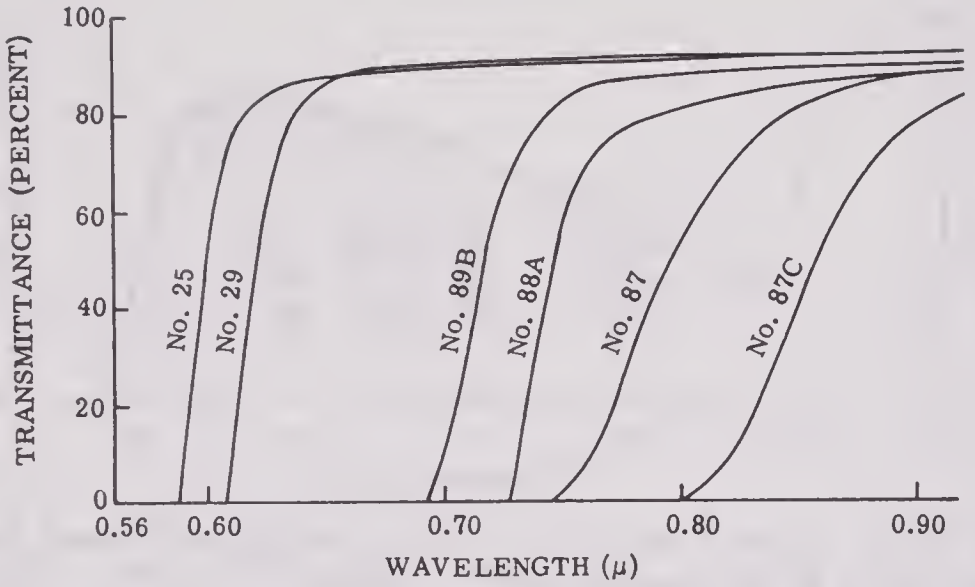


FIG. 13-15. Transmittance of Kodak Wratten filters used for infrared photography.

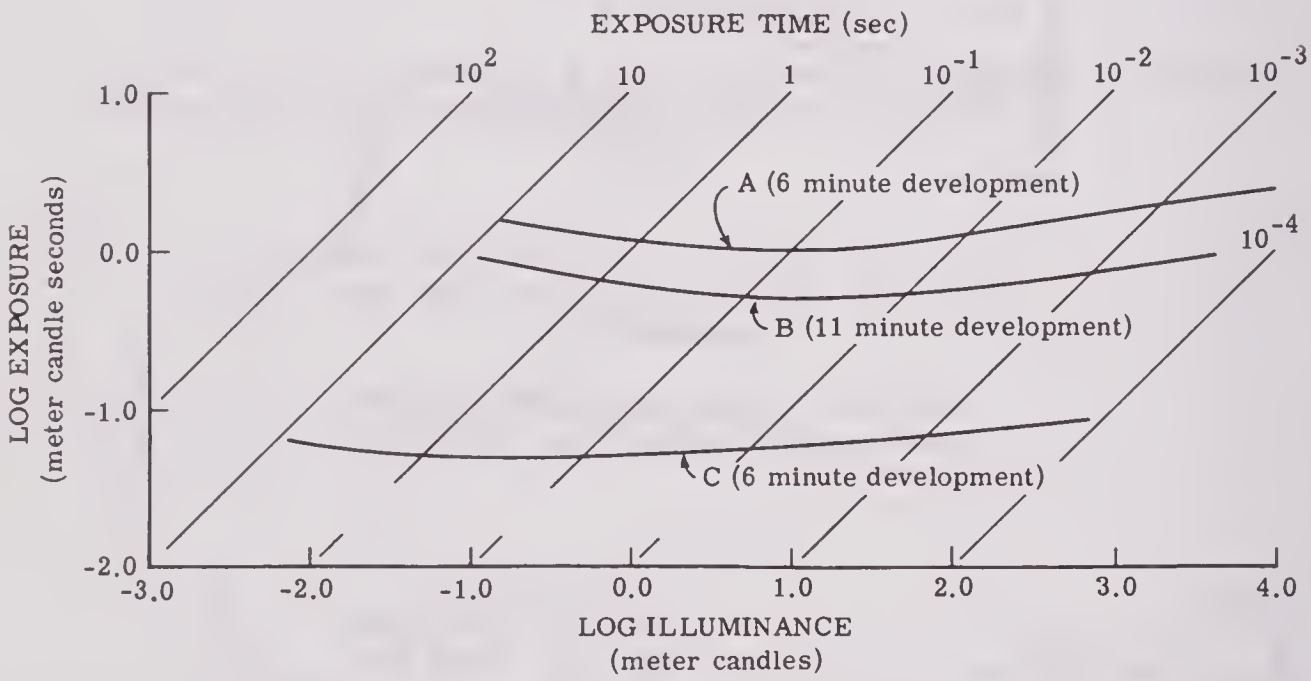


FIG. 13-16. Reciprocity characteristics for exposure to 3000°K tungsten filament through Kodak Wratten filter No. 25. Films developed in Kodak Developer D-76 at 68°F (20°C) for $D = 1.0$ above gross fog. *A*: Kodak Infrared Film. *B*: Eastman Infrared Film, Type 5210. *C*: Kodak High-Speed Infrared Film.

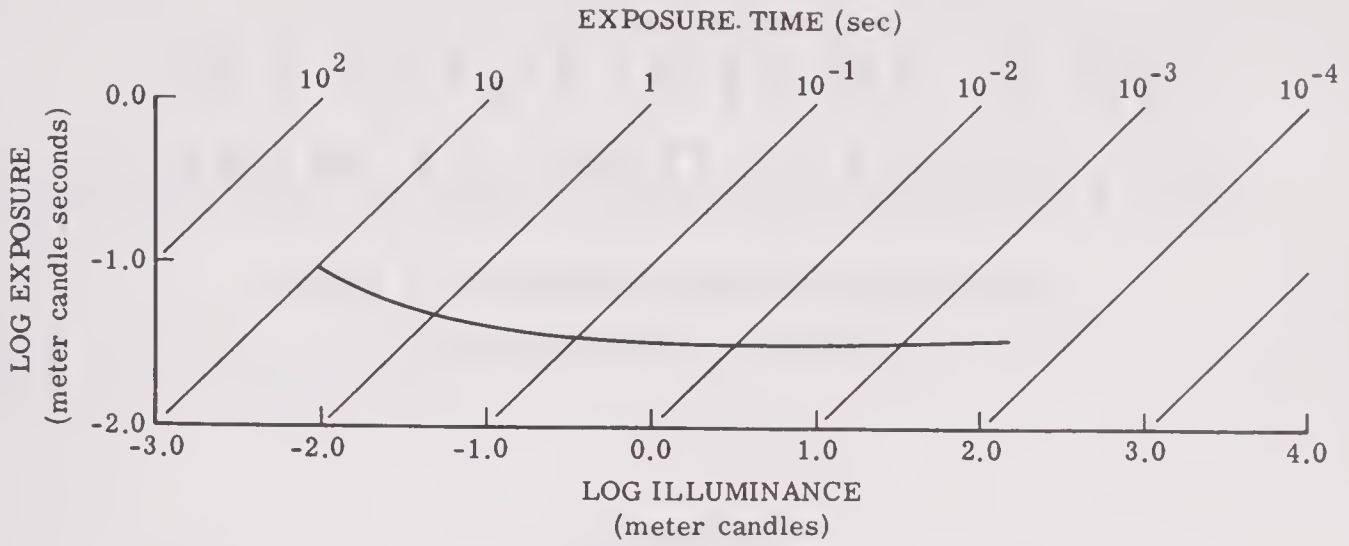


FIG. 13-17. Reciprocity characteristics of Kodak Infrared Aerographic Film for $D = 1.0$ above gross fog.

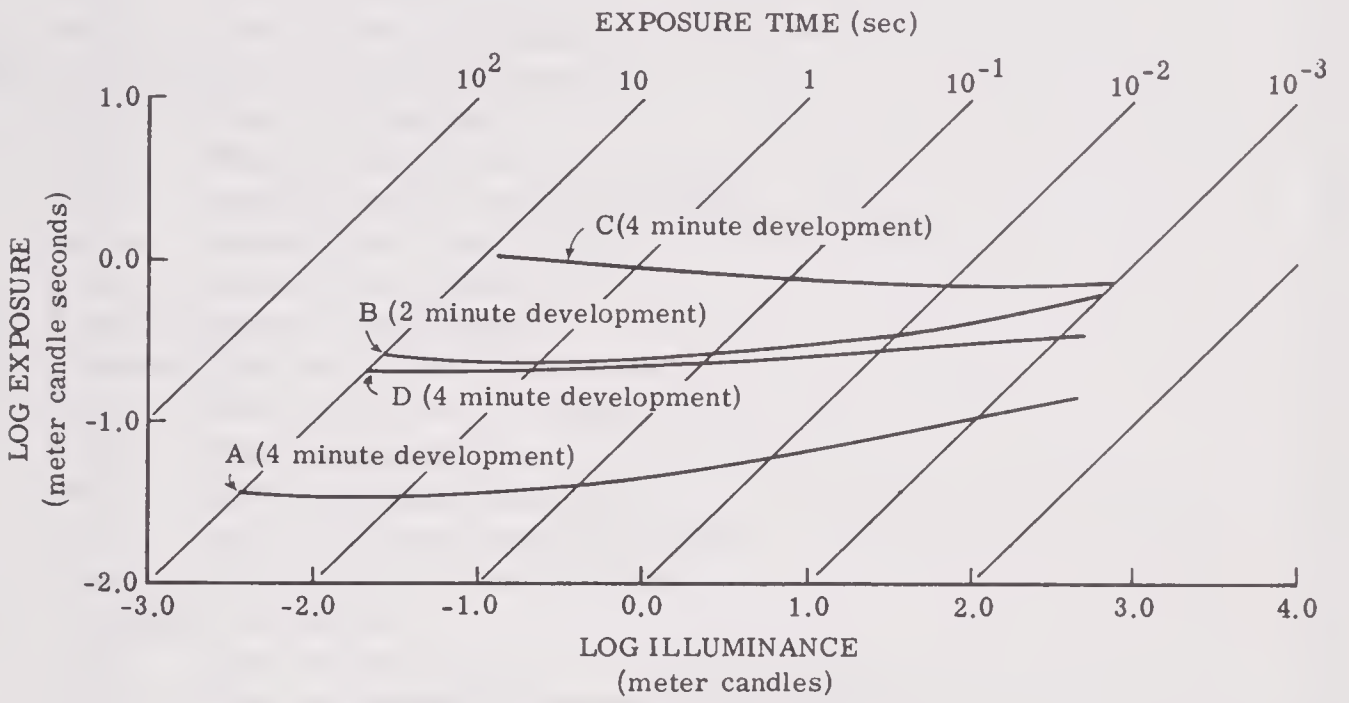


FIG. 13-18. Reciprocity characteristics for exposure to 3000°K tungsten filament through Kodak Wratten filter No. 25. Films and plates developed in Kodak Developer D-19 at 68°F (20°C). $D = 1.0$ above gross fog. Kodak Spectroscopic Plates and Films. *A*: Type I-N, water hypersensitized. *B*: Type IV-N, ammonia hypersensitized. *C*: Type I-Z, ammonia hypersensitized. *D*: Type I-M, ammonia hypersensitized.

Chapter 14

PREAMPLIFIERS AND ASSOCIATED CIRCUITS

Arthur E. Woodward and David Silvermetz

Servo Corporation of America

CONTENTS

14.1. General Requirements 584
14.2. Sources of Amplifier Noise 584
14.2.1. Thermal Noise 584
14.2.2. Shot Noise 585
14.2.3. Flicker Current Noise 586
14.2.4. Noise in Semiconductors 586
14.3. Noise Factors and Noise Figures 587
14.3.1. General Relations 587
14.3.2. Detector Noise Factor 588
14.3.3. Loading Resistance Noise Factor 589
14.3.4. Preamplifier Noise Factor 589
14.3.5. Overall Noise Factor 590
14.4. High-Frequency Compensation 590
14.5. Vacuum Tube Amplifiers 592
14.5.1. Selection of Vacuum Tubes 592
14.5.2. Input Impedance 593
14.5.3. High-Frequency Response 595
14.5.4. Very-High-Impedance Amplifier 596
14.5.5. Power Supply Considerations 597
14.6. Transistor Amplifiers 597
14.6.1. Noise *vs.* Frequency 597
14.6.2. Noise *vs.* Source Impedance and Bias Point 598
14.6.3. Transistor Bias Stabilization 598
14.6.4. Characterization of Transistor Noise 599
14.6.5. Relationship of Minimum Noise Factor to Transistor Parameters 602
14.6.6. Circuit Considerations 602
14.6.7. Transformer Coupling 604
14.6.8. Simple Low-Noise Preamplifier Design 605
14.7. Field Effect Transistors 606
14.7.1. Noise Figure 606
14.7.2. FET Preamplifier Design 608
14.8. Synchronous Detection 610
14.9. Grounding Considerations 611
14.9.1. Grounding Techniques 611
14.9.2. Low-Noise Cable 611

14. Preamplifiers and Associated Circuits

14.1. General Requirements

The characteristics of detectors used in infrared systems are detailed in Chapter 11. The impedance of these detectors extends from a few ohms to hundreds of megohms, and their output is usually in the order of microvolts.

The primary objective in preamplifier design is to produce an amplifier that will increase the detector signal to a level capable of being transmitted over a cable, possibly exceeding 10 feet in length, without degrading the signal-to-noise ratio available at the detector. The primary requirements of such an amplifier are low noise, high gain, low output impedance, large dynamic range, good linearity, and relative freedom from microphonics. The amplifier must be compact, since it is usually mounted near the detector, and must be carefully shielded to prevent the introduction of unwanted signals by stray fields.

14.2. Sources of Amplifier Noise [1]

14.2.1. Thermal Noise [2, 3]. The noise spectrum of the mean-square open circuit voltage of thermal noise is

$$S_e(f) = 4kTR \tag{14-1}$$

where k = Boltzmann's constant

T = absolute temperature (usually °K)

R = resistive impedance

The mean-square open-circuit voltage then is

$$\overline{e^2} = 4kTR\Delta f \tag{14-2}$$

where Δf = effective bandwidth

$$\Delta f = \frac{1}{G_0}\int_0^\infty G(f)\,df \tag{14-3}$$

G_0 = maximum gain

$G(f)$ = gain as a function of frequency

The power spectrum is

$$S_p(f) = kT \tag{14-4}$$

Thus R is a white thermal noise source with kT units of power per unit bandwidth.

The mean-square noise current spectrum of a Norton source can be written

$$S_i(f) = 4kTG \tag{14-5}$$

where G = shunt conductance. Both R and G are functions of frequency:

$$R = Re\,[Z_s(f)] \qquad G = Re\,[Y_{sh}(f)] \tag{14-6}$$

where Re means "real part of"

Y_{sh} = shunt admittance

More useful expressions (for multiport functions, reciprocal networks, and other special cases) are given in Bennett [1].

14.2.2. Shot Noise [4, 5]. Shot noise is caused by the independent random emission times of the discrete electrons. The number of electrons with charge emitted in a given time interval n has a Poisson distribution. The average noise current is

$$\bar{i} = I = e\bar{n} \tag{14-7}$$

For frequencies small compared to the reciprocal of the transit time, the spectral density of the mean-square shot-noise current in a temperature-limited diode is

$$S_i(f) = \bar{n}e^2 = 2eI \tag{14-8}$$

In a space-charge-limited diode with zero emission velocity, the Langmuir-Child's equation holds:

$$j = \frac{4\epsilon_0}{9}\left(\frac{2e}{m}\right)^{1/2}\frac{V^{3/2}}{d^2} \tag{14-9}$$

where j = current density

V = anode voltage

d = distance from cathode to anode

ϵ_0 = free-space permittivity (dielectric constant)

m = mass of electron

The Maxwell probability law of emission velocities is probably more valid:

$$P(v_0) = \begin{cases} (mv_0/kT_c)\exp(-mv_0^2/2kT_c) & 0 \leqslant v_0 \\ 0 \qquad \text{otherwise} \end{cases} \tag{14-10}$$

where $P(v_0)$ = the probability density that an electron will have a velocity v_0

T_c = cathode temperature

The approximate equivalent of Child's law, where $Ve \geqslant kT_e$, is

$$j = (4\epsilon_0/9)(2e/m)^{1/2}[(V_a - V_m)^{3/2}/(d - X_m)^2] \tag{14-11}$$

where V_a = potential at the anode

V_m = minimum potential

The space-charge limited diode with distributed electron velocities behaves as though its cathode were at X_m, the position of the minimum potential. Then the spectral density of anode current fluctuations is written

$$Si(f) = 2ei\gamma^2 \tag{14-12}$$

where γ^2 = the space-charge-smoothing factors. Values for γ are given by Thompson [6]. Another useful approximation is

$$i(f) = 2\ kTg = 2\ kT_{eff}\, g_d \tag{14-13}$$

where T_{eff} can be obtained from Thompson [6]

g_d is the dynamic conductance of the tube

g is an effective thermal conductance

A good approximation is

$$T_{eff} = 0.644\ T_c \tag{14-14}$$

A space-charge-limited diode can be treated as a conductance g_d operating at 0.644 times the cathode temperature.
For the triode,

$$s(f) = 2k\left(\frac{0.644\ T}{\sigma}\right)g_m \tag{14-15}$$

where g_m = dynamic transconductance

$$\sigma = \left[1 + \frac{1}{\mu}\left(\frac{d_{cp}}{d_{cg}}\right)^{4/3}\right]^{-1} \tag{14-16}$$

d_{cp} = distance from cathode to plate

d_{cg} = distance from cathode to grid

For the pentode,

$$S(f) = (\bar{i}_a\bar{i}_2/\bar{i}_c^2)e\bar{i}_c + (\bar{i}_a/\bar{i}_c)^2\ [e\bar{i}_c\gamma^2 + \bar{i}_c^2\delta(f)] \tag{14-17}$$

i_a = anode current

i_2 = second grid current

i_c = cathode current

$\delta(f)$ = impulse at zero frequency

γ^2 = smoothing factor

Additional results for multigrid and negative grid tubes, including the effects of flicker and partition noise, are given in Bennett [1] in Chapter 4.

14.2.3. Flicker Current Noise. The mechanism of flicker noise ($1/f$ noise) has not been completely clarified. It is characterized by a power spectrum that varies inversely with frequency, hence the name $1/f$ noise. In vacuum tubes, at low and sub-audio frequencies, flicker noise can exceed the shot noise by several orders of magnitude. If flicker noise is represented by a noise-current generator, then the noise spectrum is

$$\bar{i}^2 = K\bar{i}^2\int_{f_1}^{f_2} 1/f\,df \tag{14-18}$$

where K is a constant. In many semiconductors, flicker noise is also the dominant source of noise at low frequencies. The noise results from a fluctuation in current density caused by some modulation mechanism, and can be divided into two components. One component causes random changes at the transistor surface; the other causes random changes in the leakage path around the transistor junctions (primarily the reverse-biased collector junction).

14.2.4. Noise in Semiconductors. The mean square current noise spectrum can be written

$$S_i(f) = 4kTG(f) + 2eI \tag{14-19}$$

The shot noise of a junction transistor can be represented by noise generators as shown in Fig. 14-1.

$$\overline{v_{bn}^2} = 4kTz_b \tag{14-20}$$

$$\overline{i_{en}^2} = 4kTy_e(f) - 2eI_e \tag{14-21}$$

$$v_{1n} = -i_{en}z_e$$

$$\overline{i_{2n}^2} = i_{cn} - \alpha i_{en} = 2e\alpha(1-\alpha)I_e\Delta f \tag{14-22}$$

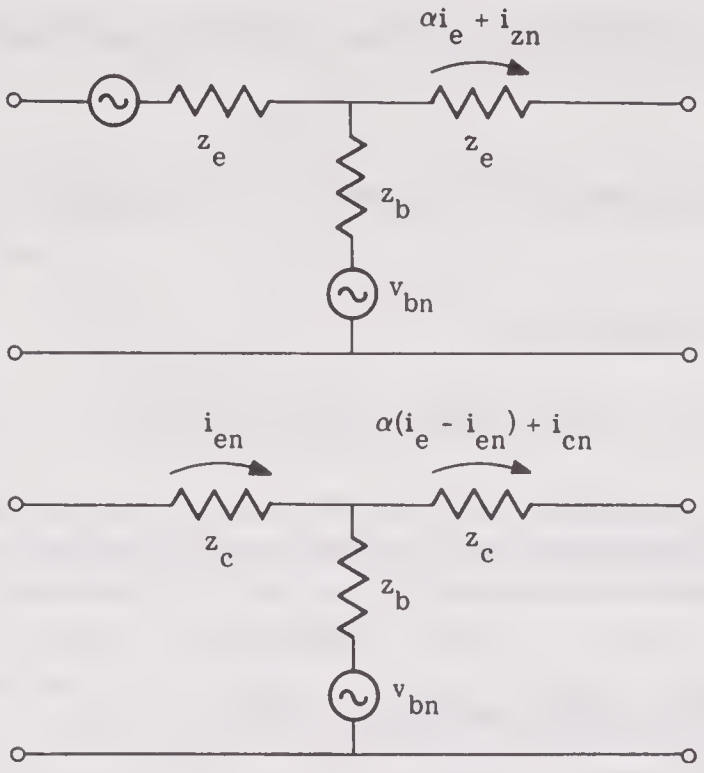


FIG. 14-1. Noise circuit for transistor.

Generation-recombination (g-r) noise [7] is a bulk property of semiconductor material and results from fluctuations in the conductivity produced by carrier-density changes. The variation in carrier-density is caused by the random character of generation, recombination, and trapping processes. Depending on the magnitude of the current noise, g-r noise may be impossible to detect. As a result, the semiconductor exhibits current noise at low frequencies and thermal noise at high frequencies. The magnitude of g-r noise is given by

$$\overline{i_{g\text{-}r}^2} = \frac{2I^2\tau\Delta f}{N\,[1 + (2\pi f\tau)^2]} \tag{14-23}$$

where τ = electron-hole lifetime

N = average total number of free electrons

f = frequency of modulating signal

14.3. Noise Factors and Noise Figures [5, 8]

14.3.1. General Relations. The available (incremental) power ΔP_{as} of a source P_{as} is the maximum average power obtainable (see Fig. 14-2):

$$\Delta P_{as} = \overline{e_s^2}/4R_s = kT_s\Delta f \tag{14-24}$$

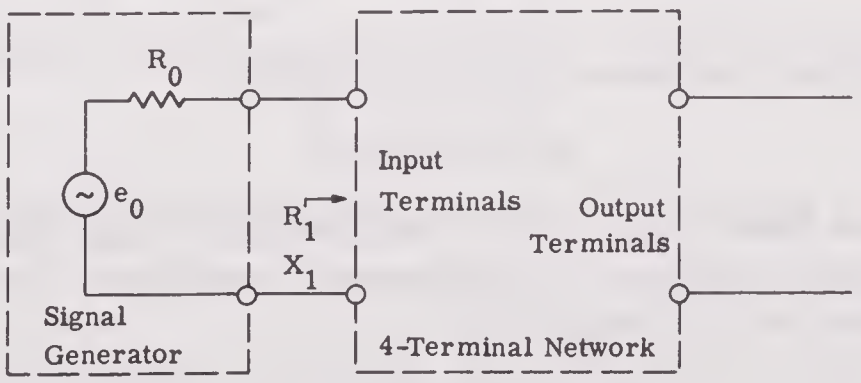


FIG. 14-2. Noise relations in a two-terminal pair network.

The effective noise temperature T_{es} is the available noise from the source N_{as} divided by $k\Delta f$:

$$T_{es} = N_{as}/k\Delta f \tag{14-25}$$

The standard noise temperature T_0 is 290°K and the relative noise temperature t_s is given by

$$t_s = T_{es}/T_0 \tag{14-26}$$

The available power gain G_a is:

$$G_a = P_{ao}/P_{as} = \frac{R_s}{R_o}\left|\frac{H(\omega)Z_i}{Z_i + Z_s}\right|^2 \tag{14-27}$$

The operating noise factor F_0 is the total available noise output power N_{a0} divided by the available noise output power from the source alone N_{a0_s}:

$$F_0 = N_{a0}/N_{a0_s} \tag{14-28}$$

The noise power from an ideal generator is $kT\Delta f$.

Other expressions are

$$F_0 = \frac{(S/N)_{as}}{(S/N)_{a0}} \tag{14-29}$$

$$F_0 = 1 + \frac{N_{a0}}{N_{a0_s}} = 1 + \frac{N_{a0_i}}{G_a k T_{es} \Delta f} \tag{14-30}$$

$$F_0 = 1 + \frac{T_0}{T_{es}}(F - 1) \tag{14-31}$$

$$F = 1 + \frac{N_{a0_i}}{G_a k T_0 \Delta f} \tag{14-32}$$

where N_{a0_i} = internally generated noise power

S = signal power

For cascaded stages 1 .... N

$$G_a\Big|_1^N = G_a\Big|_1^{N-1} G_a\Big|_N = \prod_{n=1}^{N} G_{a_n} \tag{14-33}$$

$$F\Big|_1^N = F_1 + \sum_{n=2}^{N}\left[(F_n - 1)\Big/\prod_{m=1}^{n-1} G_{a_m}\right] \tag{14-34}$$

$$F_0\Big|_1^N = F_{01} + \sum_{n=2}^{N}\left[(F_{0n} - 1)\Big/\prod_{m=1}^{n-1} G_{a_m}\right] \tag{14-35}$$

The noise figure is the noise factor expressed in decibels,

$$NF = 10\log_{10} F \tag{14-36}$$

14.3.2. Detector Noise Factor. The noise produced at the output of an infrared detector is discussed in Chapter 11. The detector's noise factor, which is many times greater than thermal noise, is found to be

$$F_D = \alpha k T \Delta f / k T \Delta f = \alpha \tag{14-37}$$

where α is the excess detector noise.

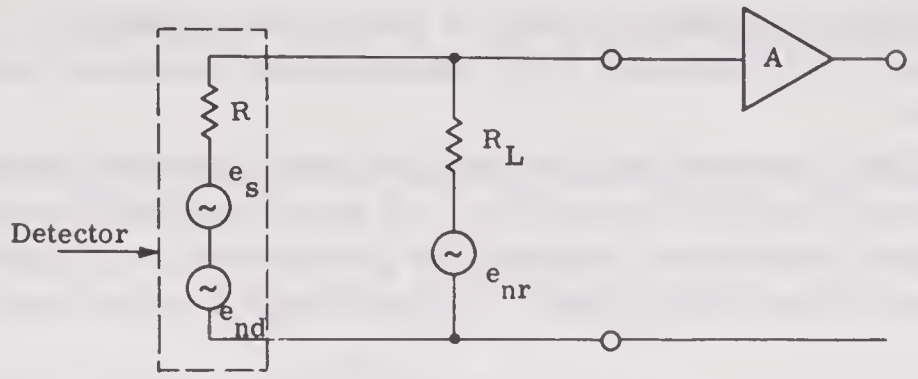


FIG. 14-3. Loading resistance noise factor network.

14.3.3. Loading Resistance Noise Factor. In the network shown in Fig. 14-3 loading resistance R_L represents all resistance elements shunting the detector. This resistance includes all biasing resistances and the input resistance to the preamplifier. The detector is represented by a noiseless resistor in series with a signal generator and a noise generator. The loading resistance is represented by a noiseless resistor in series with the noise generator alone. The available power gain is given by

$$G_a = G = \frac{S}{S_g} = \frac{e_s^2 \left(\frac{R_L}{R+R_L}\right)^2 \Big/ \frac{4RR_L}{R+R_L}}{e_s^2/4R} \tag{14-38}$$

which reduces to

$$G = \frac{R_L}{R+R_L} \approx 1 - \frac{R}{R_L} + \left(\frac{R}{R_L}\right)^2 - \ldots + \ldots \tag{14-39}$$

The noise factor from R_L is then

$$F_L = \frac{N}{GkT\Delta f} = \frac{kT\Delta f}{(R_L/R+R_L)\,kT\Delta f} = 1 + \frac{R}{R_L} \tag{14-40}$$

Thus, R_L should be as large as practical. However, a large R_L introduces additional problems in the form of shunting capacitance.

The noise factor of the detector in cascade with loading resistance is given by

$$F_{DL} = F_D + \frac{F_L - 1}{G_D} = \alpha + \frac{R}{R_L} \tag{14-41}$$

14.3.4. Preamplifier Noise Factor. At low frequencies, vacuum tube noise can be represented by an emf in series with the grid having the relation

$$emf = (4kTR_{eq}\Delta f)^{1/2} \tag{14-42}$$

where R_{eq} = equivalent noise resistance of the tube. This can be written as [9]

$$R_{eq} = R_s + b/f \tag{14-43}$$

The first term of this equation is the shot noise R_s. For a triode,

$$R_s = 2.5/g_m \tag{14-44}$$

where g_m = transconductance of the tube. The second term b/f of Eq. (14-43) is $1/f$ noise (flicker noise). For a good tube ($I_0 = 1$ ma), the factor b is of the order of 10^6; for average tubes, b is about 10^7; and for poor tubes it is about 10^8 or 10^9. The factor b

is current-dependent and passes through a minimum of about $I_0 = 1$ ma. Vacuum tubes should therefore be operated below recommended operating conditions for low-noise applications.

In transistors, two noise sources (generally partially correlated) are used to describe the equivalent noise input to the amplifier. (A more complete discussion is given in Sec. 14.6.) Figure 14-4 shows a preamplifier represented by a noiseless amplifier in series with an equivalent noise resistor. The preamplifier noise factor is then

$$F_A = 1 + R_{eq}\left[\frac{R + R_L}{RR_L}\right] \tag{14-45}$$

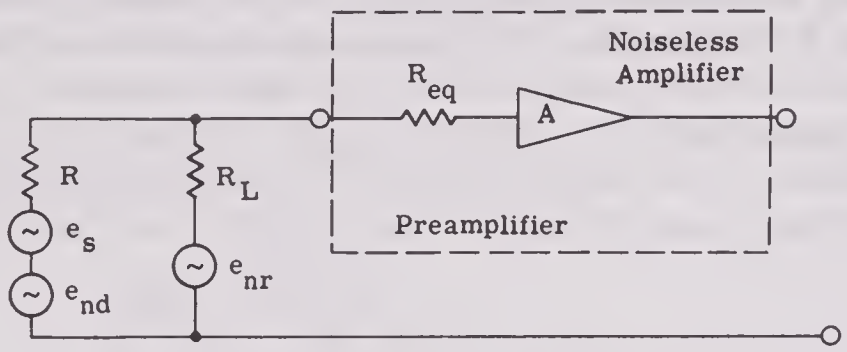


FIG. 14-4. Preamplifier noise factor network.

14.3.5. Overall Noise Factor. The overall noise factor (including detector noise, loading resistance noise, and preamplifier noise) is

$$F_{DA} = \alpha + \frac{R}{R_L} + R_{eq}R\left(\frac{R + R_L}{RR_L}\right)^2 \tag{14-46}$$

The overall signal-to-noise ratio is

$$\frac{S_0}{N_0} = \frac{1}{F_{DA}}\frac{S_i}{N_i} \tag{14-47}$$

14.4. High-Frequency Compensation

To minimize the loss in the available signal-to-noise ratio of the detector, the resistance in parallel with the detector should be as large as possible. A high source impedance requires high-frequency compensation in an amplifier to correct for bandwidth loss due to capacitive input loading. Figure 14-5 illustrates the actual input circuit at high frequencies where R = Thevenin equivalent of the detector and parallel resistances and C = all capacitive loading at amplifier input.

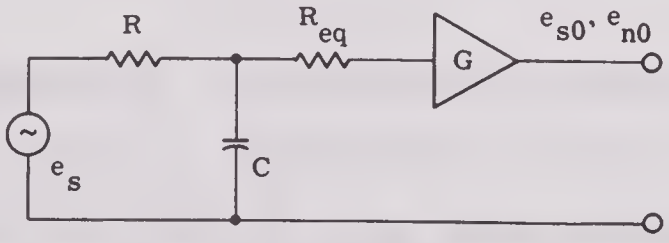


FIG. 14-5. Preamplifier input circuit at high frequencies.

The magnitude of the signal voltage at the grid of the preamplifier is

$$e_{sg} = e_s/(1 + \omega^2 R^2 C^2)^{1/2} \tag{14-48}$$

The mean-square noise voltage at the grid of the preamplifier due to the thermal noise of R is

$$\overline{e_{ng}^2} = 4kT \int_{f_1}^{f_2} \frac{R}{1 + \omega^2 R^2 C^2}\, df \tag{14-49}$$

If the amplifier has a response such that

$$G = G_0 (1 + \omega^2 C^2 R^2) \tag{14-50}$$

then

$$e_{s0} = (G_0)^{1/2} e_s \tag{14-51}$$

and the noise voltage at the output is given by

$$\overline{e_{n0}^2} = 4kT \int_0^{f_2} RG/(1 + \omega^2 R^2 C^2)\, df = 4kTRG_0 \Delta f \tag{14-52}$$

Neglecting the shot noise of the first stage, the signal-to-noise ratio at the output is

$$e_{s0}/e_{n0} = e_s/(4kTR\Delta f)^{1/2} \tag{14-53}$$

This is identical to the signal-to-noise ratio obtained at the input in the absence of any shunt capacitance. Therefore, by using a properly compensated amplifier where $G = G_0(1 + \omega^2 R^2 C^2)$ the effect of input capacitance is eliminated without affecting the signal-to-noise ratio.

The noise voltage from R_{eq} at the grid of the preamplifier input tube is given by

$$e_{nr} = (4kTR_{eq}\Delta f)^{1/2} \tag{14-54}$$

The noise voltage at the output of the compensated amplifier is called peaked-channel noise. It is found to be

$$\overline{e_{nr0}^2} = 4kTR_{eq}G_0 \int_{f_1}^{f_2} (1 + \omega^2 R^2 C^2)\, df \tag{14-55}$$

which becomes

$$\overline{e_{nr0}^2} = 4\pi kTR_{eq}G_0 [\Delta f + 4\pi R^2 C^2 (\Delta f)^3/3] \tag{14-56}$$

where $\Delta f = f_2 - f_1$. Peaked-channel noise reduces the fine detail of a signal, but it can be tolerated better than flat-channel or white noise which, because of its broad spectral structure, could completely distort the signal.

In many instances, the shot noise can be neglected. If the upper limit of the input circuit of a dc system is given by

$$f_2 = 1/2\pi RC \tag{14-57}$$

and the amount of compensation is designated as

$$m = f_2/f_1 = \Delta f/f_1 \tag{14-58}$$

then

$$e_{nr_0} = (G_0)^{1/2} (4kTR_{eq}\Delta f)^{1/2} (1 + m^2/3)^{1/2} \tag{14-59}$$

The total ouput noise due to thermal noise R and the shot noise of the input tube is given by

$$e_{nt} = (\overline{e_{n0}^2} + \overline{e_{nr0}^2})^{1/2} \tag{14-60}$$

which is the same as

$$e_{nt} = (4kT\Delta f G_0)^{1/2} [R_{eq}(1 + m^2/3) + R]^{1/2} \tag{14-61}$$

This is shown in the following example where

$$R_{eq} = 1 \text{ k}\Omega \text{ (triode)}$$
$$R = 200 \text{ k}\Omega$$
$$C = 4 \text{ pf}$$
$$f_3 = 500 \text{ kc}$$

This produces the relationships

$$f_1 = 0 \tag{14-62}$$

$$f_2 = (2\pi RC)^{-1} = [2\pi(2 \times 10^5)(4 \times 10^{-12})]^{-1} = 200 \text{ kc} \tag{14-63}$$

$$m = f_3/f_2 = 500/200 = 2.5 \tag{14-64}$$

$$R_{eq}(1 + m^2/3) = 1000[1 + (2.5)^2/3] = 3080 \tag{14-65}$$

Comparing this value with the value of R produces

$$R >> R_{eq}(1 + m^2/3) \tag{14-66}$$

The shot noise can therefore be neglected. The noise voltage then reduces to

$$e_{nr_0} = (4kTR\Delta f G_0)^{1/2} \tag{14-67}$$

which is what is obtained when shot noise is neglected.

14.5. Vacuum Tube Amplifiers

Low-noise vacuum tubes are useful as preamplifier stages for high-impedance infrared detectors because they provide relatively high gain and wide bandwidths with high input impedances. Vacuum tube amplifiers have good dynamic range and agc characteristics. Their disadvantages are microphonics, high power requirements, and, before the development of the Nuvistor and ceramic tubes, their relatively large size. Microphonics can be greatly minimized by using low-microphonic tubes and judicious mounting techniques. The power requirements, however, are still much greater than those of the semiconductors.

The preamplifier must be mounted near the detector; the number of stages used depends upon the gain required, the bandwidth, and the space available. Since most detectors are high-impedance sources, they act as constant-current devices whose signal-voltage output is a function of the loading resistance. This resistance should be high enough that the thermal noise generated by the loading resistance does not limit the signal-to-noise ratio available from the detector. In order to prevent a loss in high-frequency response, input capacitance effects must be compensated for.

The design of several preamplifiers used with various detectors is discussed in the following paragraphs. They illustrate practical solutions to specific problems and do not necessarily represent the best amplifier circuits available.

14.5.1. Selection of Vacuum Tubes. The major considerations in the selection of vacuum tubes for preamplifiers are low noise, low microphonic response, and available space. The shot noise of a triode is represented by the equivalent noise resistance given in Eq. (14-44). In the pentode, the shot noise can be represented by

$$R_s = \frac{I_b}{I_b + I_{g2}} \left(\frac{2.5}{g_m} + \frac{20\, I_{g2}}{g_m{}^2} \right) \tag{14-68}$$

where I_b = average plate current

I_g = average screen current

Since the shot noise of a pentode is considerably higher than that of a triode, triodes should be used in preamplifiers. The calculated equivalent noise resistances of various tubes are given in Table 14-1. Vacuum tubes used for low-noise, low-microphonic preamplifiers are:

6112 dual triode
CK6533WA single triode
CK8096 single triode
12AY7 dual triode

TABLE 14-1. EQUIVALENT NOISE RESISTANCES OF VARIOUS TUBES

Tube Type	g_m	R_{eq}
6AK5 (triode connected)	6,670	375
6F4 triode	5,800	430
417A triode	24,000	90
6J4 triode	12,000	210
6AK5 pentode	5,000	1,880
6AG5 pentode	5,000	1,640
436A pentode	28,000	210
6112 triode	2,500	1,000
CK8096 triode	1,750	1,400
6CW4 triode	12,500	200

14.5.2. Input Impedance. High input impedance in vacuum tube amplifiers can be obtained in two ways. The first is to utilize a large grid resistor. Values up to 10 MΩ can be used, while values limited to 1 MΩ or less are not uncommon. The second is the use of negative feedback. This method increases the input impedance, stabilizes amplifier gain, and improves the low-frequency response. In addition, the use of negative feedback does not affect the signal-to-noise ratio at the output of the amplifier [10].

The use of negative feedback in an amplifier is shown in Fig. 14-6. The closed-loop gain is given by

$$K = \frac{A}{1 + A\beta} \tag{14-69}$$

where A = open-loop gain of the amplifier.

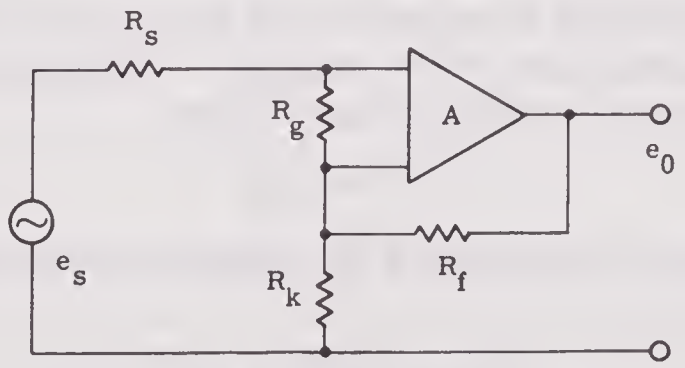


FIG. 14-6. Negative feedback in amplifiers.

$$\beta = R_K/R_f, \qquad \text{for } R_f >> R_K \tag{14-70}$$

The input impedance is given by

$$R_{in} = r_g\,(A/K) \tag{14-71}$$

The input impedance is determined by the largest grid resistor that can be used and the open loop gain available. A loop gain $A\beta$ of 4 or greater is required to provide adequate gain stabilization and a useful increase in input resistance. Figure 14-7 shows a single dual triode used as a low-noise preamplifier and having a bandwidth of 1 to 1000 cps. The amplifier constants are

$$A = 1500$$
$$R_g = 2\ \text{M}\Omega$$
$$R_f = 300\ \text{k}\Omega$$
$$R_k = 390\ \Omega$$
$$K = 300$$
$$R_{in} = 10\ \text{M}\Omega$$

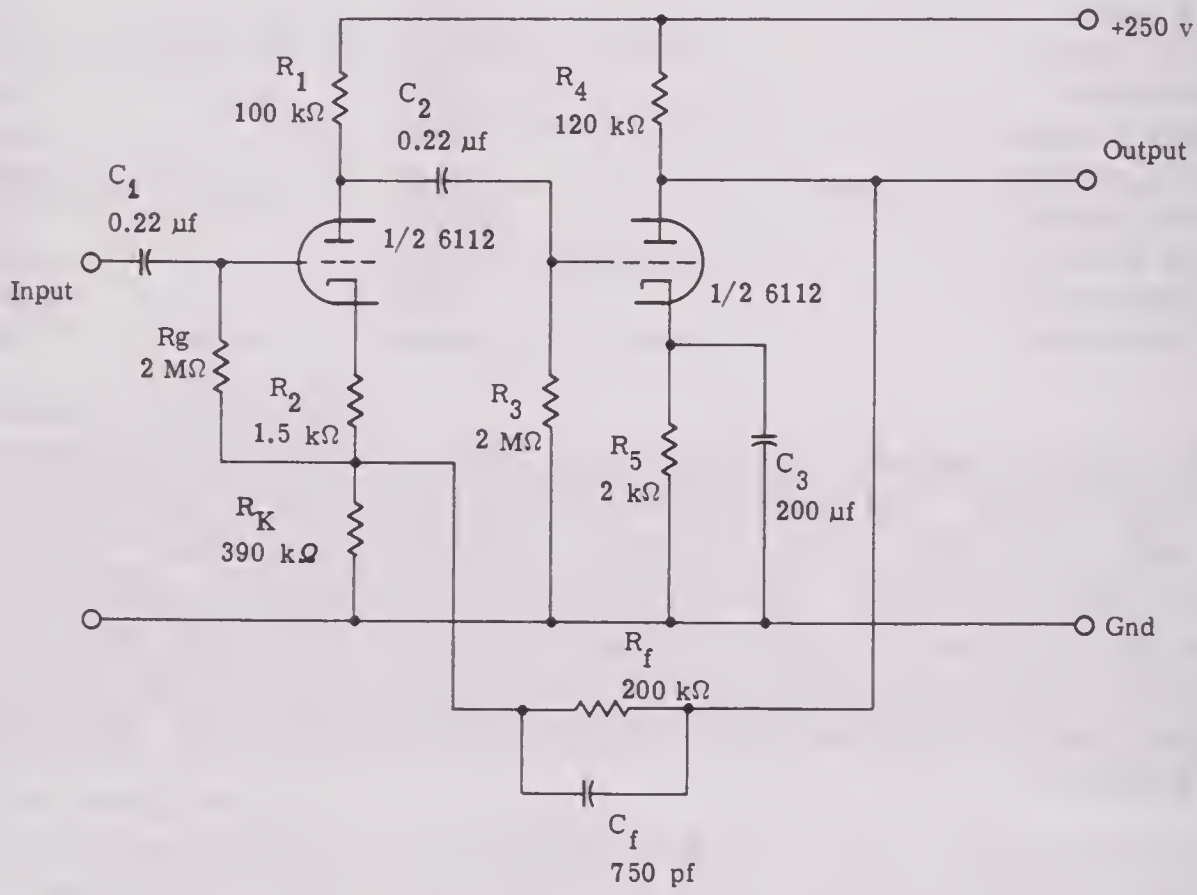


FIG. 14-7. Low-noise preamplifier.

Figure 14-8 shows the effect of feedback on the signal-to-noise ratio. A noise voltage e_j is injected between gain stages of the amplifier. Reflecting this voltage back to the signal source as an equivalent noise voltage e'_j gives

$$e'_j = e_j/G_1 \tag{14-72}$$

The signal-to-noise ratio at the output of the amplifier without feedback (switch S open) is then

$$\frac{e_{s0}}{e_{n0}} = \frac{G_1 G_2 e_s}{G_1 G_2 e'_j} = G_1 \frac{e_s}{e_j} \tag{14-73}$$

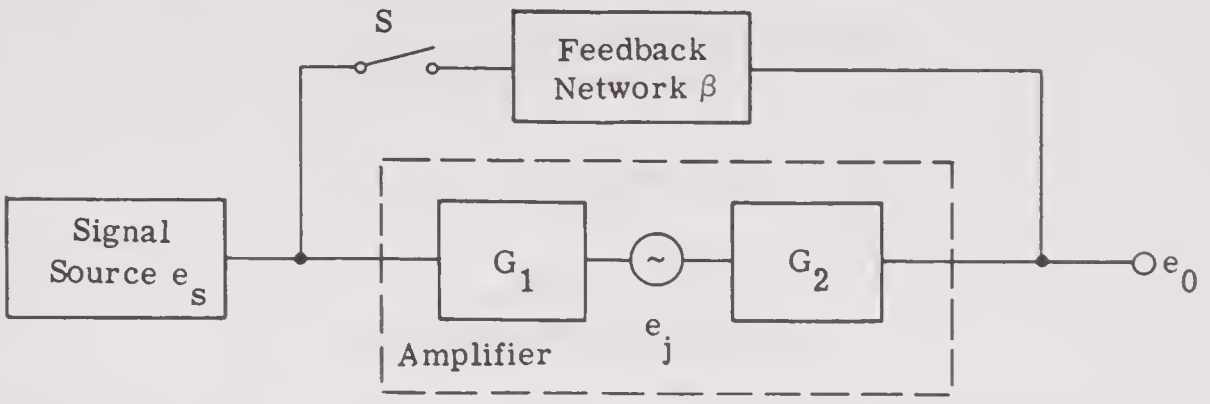


FIG. 14-8. Effect of feedback on noise.

The signal-to-noise ratio at the output of the amplifier with feedback (switch S closed) is given by

$$\frac{e'_{s0}}{e'_{n0}} = \frac{Ae_s/(1 - A\beta)}{Ae_j/(1 - A\beta)} = \frac{G_1 e_s}{e_j} \tag{14-74}$$

Therefore, the signal-to-noise ratio is independent of the loop gain A, regardless of the noise source.

14.5.3. High-Frequency Response. Wide bandwidths and low detector resistances are essential in infrared systems. A low-noise, low-input-capacitance amplifier is required in order to obtain optimum performance from the newer detectors. Triode amplifiers are seriously limited by the Miller effect which increases the input capacitance by the gain of the stage. Pentodes have low input capacitance, but the screen-plate partition noise adds to the problem. The cascode amplifier (see Fig. 14-9) provides the low noise characteristic of a triode combined with the gain and stability of a pentode. It consists of a grounded-cathode amplifier with the input impedance of a grounded-grid stage acting as the plate load. The gain of the grounded-cathode first stage is given by [11].

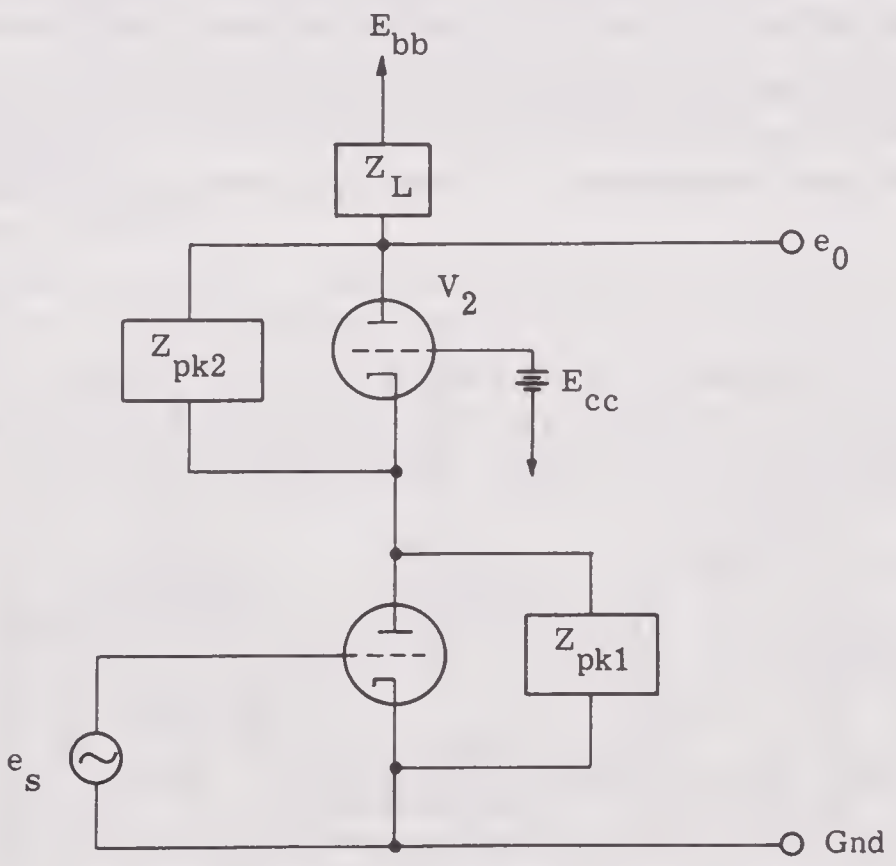


FIG. 14-9. Cascode amplifier.

$$G_1 = \frac{-\mu_1 Z_{pk1} Z_{i2}/(Z_{pk1} + Z_{i2})}{r_{p1} + Z_{pk1}\, Z_{i2}/(Z_{i2} + Z_{pk1})} \tag{14-75}$$

where

$$Z_{i2} = \frac{r'_{p2} + Z_L}{\mu'_2 + 1} \tag{14-76}$$

$$r'_{p2} = \frac{r_{p2}}{1 + j\omega C_{pk2} r_{p2}} \tag{14-77}$$

$$\mu'_2 = \frac{\mu_2}{1 + j\omega C_{pk2} r_{p2}} \tag{14-78}$$

The gain of the grounded-grid second stage is given by

$$G_2 = \frac{(\mu'_2 + 1)\, Z_L}{r'_{p2} + Z_L} \tag{14-79}$$

The overall gain is

$$G = G_1 G_2 \tag{14-80}$$

At low frequencies, where $Z_L = R_L$ and the interelectrode impedances can be neglected, the gain becomes

$$G = \frac{-\mu_1(\mu_2 + 1)R_L}{r_{p1}(\mu_2 + 1) + r_{p2} + R_L} \tag{14-81}$$

Furthermore, if $r_{p1}(\mu_2 + 1) >> r_{p2} + R_L$, then the gain reduces to

$$G \approx -g_{m_1} R_L \tag{14-82}$$

The cascode amplifier therefore has the gain of a pentode.

The gain of the first stage is low as a result of the low effective plate load. Therefore, the effective input capacitance is also low. Two problems associated with the use of this circuit are (1) the need to provide a fixed bias for the grid of the grounded-grid stage, and (2) the need for a high-plate-voltage source to supply the series-connected tubes. The noise properties of this amplifier are discussed in [12].

14.5.4. Very-High-Impedance Amplifier. The very-high-impedance amplifier shown in Fig. 14-10 uses a cascode amplifier with feedback to the input grid to further reduce the effect of input capacitance. It is used as a summing or operational amplifier

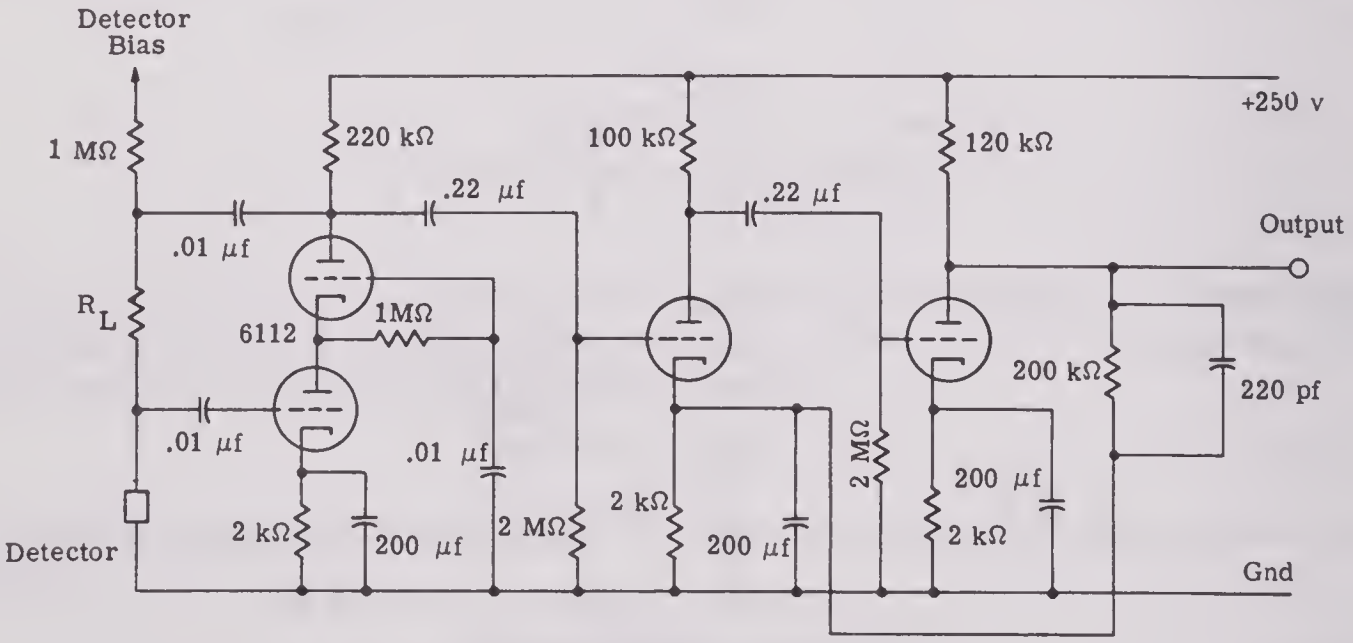


FIG. 14-10. Very-high-impedance amplifier.

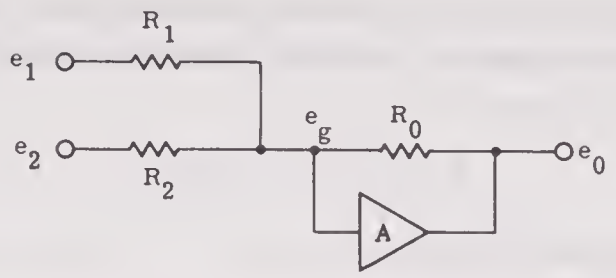


FIG. 14-11. Summing amplifier.

in analog computers. The equations for its transfer function are developed on the basis of the circuit for the summing amplifier shown in Fig. 14-11. The nodal equation for the input grid is

$$[(e_0/A)-e_1]\ R_1^{-1} + [(e_0/A)-e_2]\ R_2^{-1} + [(e_0/A-e_0]\ R_0^{-1} = 0 \qquad (14\text{-}83)$$

By simplifying and rearranging terms, where the feedback coefficient is given by

$$\beta = [1 + (R_0/R_1)\ (R_0/R_2)]^{-1} \qquad (14\text{-}84)$$

and where $A\beta$ is very large, R_1 and $R_2 >> R_0$, and e_2 is a ground potential, one obtains

$$e_0/e_1 = -R_0/R_1 \qquad (14\text{-}85)$$

Thus, when the amplifier loop gain is very high, the output is independent of the actual gain and is a gunction only of the ratio of R_0 to R_1.

The input impedance of a summing amplifier is given by

$$R_{in} = R_0/A \qquad (14\text{-}86)$$

As a result, the effects of input capacitance are substantially reduced. However, the reduction in input impedance results in a stage gain of approximately unity (if $R_0 = R_1$). Therefore, additional stages of low-noise amplification are needed to provide the required gain. (See Fig. 14-10 where the first amplifier stage is followed by a low-noise amplifier similar to the one shown in Fig. 14-7.)

14.5.5. Power Supply Considerations. Because of the very low level signals amplified by preamplifiers, the noise level introduced through the power supplies should be much less than the noise introduced at the input. In most preamplifiers, the noise or ripple voltage on the B+ supply should be less than 200 μv. For the filament supply, 500 μv is the maximum.

14.6. Transistor Amplifiers

14.6.1. Noise vs. Frequency. The power spectrum of transistor noise *vs.* frequency is given in Fig. 14-12. For typical planar transistors the break point f_1 occurs between 100 and 1000 cps. Above f_1, shot noise, which has a flat power spectrum, predominates. Below f_1, the noise figure varies inversely with frequency, with a slope of approximately 3 db per octave. This region, termed the $1/f$ region, is dominated by surface and leakage noise.

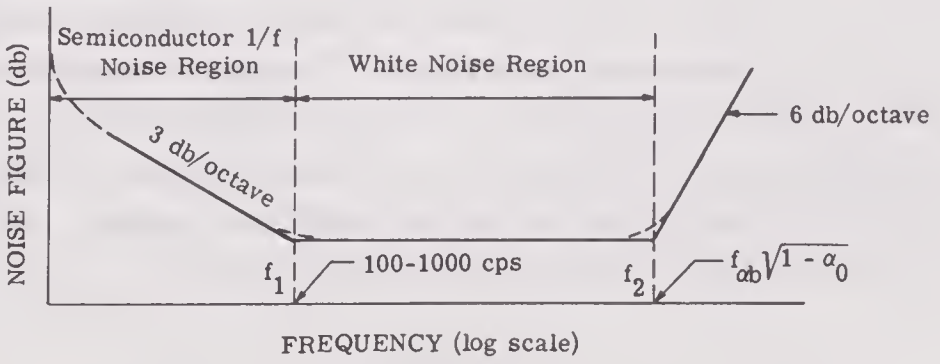


FIG. 14-12. Transistor noise *vs.* frequency power spectrum.

At a frequency f_2 the noise figure begins to increase with a slope of 6 db per octave. This increase is caused by a decrease in power gain rather than an increase in internal noise. The frequency f_2 is given by

$$f_2 \simeq f_{\alpha b} (1 - \alpha_0)^{1/2} \tag{14-87}$$

where $f_{\alpha b}$ = common base cutoff frequency

α_0 = current gain

14.6.2. Noise vs. Source Impedance and Bias Point. Figure 14-13 shows the broadband noise figure of a low-noise amplifier having a half-power passband of 10 to 10,000 cps. The noise figure is given as a function of source impedance for three different values of emitter current. Figure 14-13 shows that the optimum source impedance increases as the emitter current is decreased.

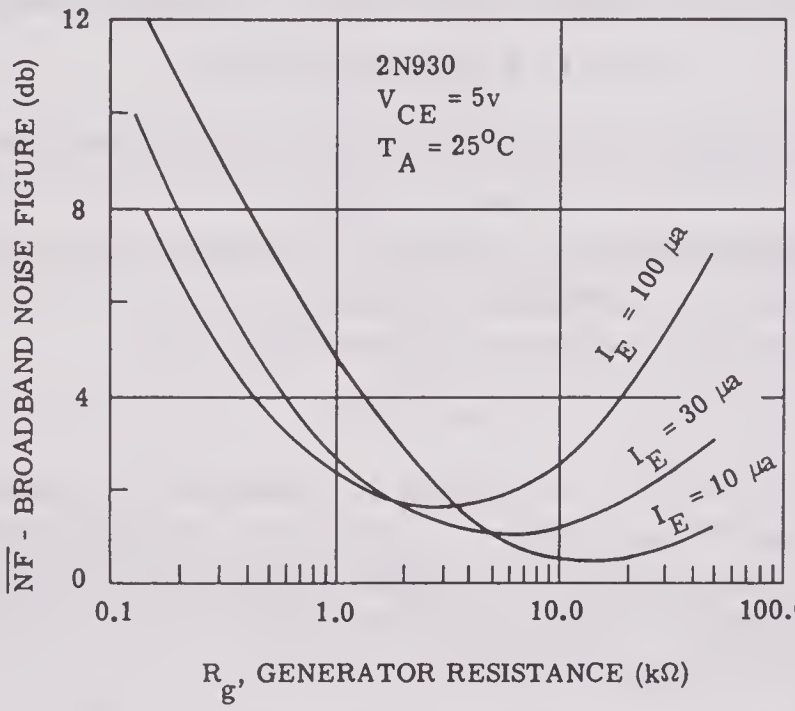


FIG. 14-13. Broadband noise figure *vs.* source impedance.

14.6.3. Transistor Bias Stabilization.

14.6.3.1. *Bias Stabilization and Dynamic Range.* Since the noise figure is a function of collector current, transistors must be biased for low-noise operation (generally between 10 μa and 200 μa). As the source impedance increases, the optimum collector current usually decreases. The use of the low-noise region is limited by the problems of maintaining bias stabilization through fluctuations in ambient temperature and changing equipment parameters. However, the biasing problem is greatly reduced by using low-noise silicon planar transistors which have low collector leakage currents and high current gains.

A transistor stage cannot accept any signal large enough to cause the collector to swing into saturation or cutoff. Drift in the dc collector-to-emitter voltage of a properly designed stage lowers this acceptable signal level, thereby lowering the dynamic range of the preamplifier. Thus, dynamic range is reduced if bias stability is reduced.

14.6.3.2. *Bias Stabilization Relating to Low-Noise Designs.* Transistor stability is the ratio of the incremental change in the total collector current, I_c, to an incremental change in the collector leakage current, I_{c0}, and is given by

$$S = \frac{\partial I_c}{\partial I_{c0}} \tag{14-88}$$

If R_B is the equivalent dc resistance in the base circuit and R_E the equivalent dc resistance in the emitter circuit, then the stability is approximately

$$S = \frac{R_B + R_E}{R_B(1 + h_{fb}) + R_E} \tag{14-89}$$

For values of h_{fb} almost equal to -1, the expression reduces to

$$S \approx \frac{R_E + R_B}{R_E} \tag{14-90}$$

In high-impedance RC-coupled amplifiers, R_B is large in order to maintain a high input impedance. If $R_B >> R_E$, the stability approaches the forward current gain of the transistor (h_{FB}). In order to improve the stability, the emitter resistance R_E must be increased. However, as R_E is increased, the allowable collector swing is decreased. Therefore, each design involves compromises between opposing factors.

Before the development of planar transistors, an I_{c0} of less than 20 na (nanoamperes) at 25° C was rare in high-frequency transistors. The I_{c0} would double for approximately every 10° C increase, so that at 85° C, $I_{c0} = 1.3$ μa. With a typical stability of 5, the change in I_c would have been 5.5 μa, making operation at a nominal collector current of 10 μa (a good low-noise point) impossible.

The typical collector leakage current of a good low-noise planar transistor is less than 1 na at 25° C, and doubles with approximately every 14° C increase. At 85° C, with I_{c0} less than 52 na, and with a bias stability of 5, the change in I_c caused by the change in I_{c0} would be less than 0.25 μa. In this case, operation at a nominal collector current of 10 μa would be feasible.

In Eq. (14-89), if R_B is 1 MΩ and $S = 5$, then R_E must be of the order of 250 kΩ (assuming h_{FB} close to unity). The power supply voltage dropped across R_E is then 10 μa times 250 kΩ, or 2.5 volts. If a stability of 3 is desired, R_E must increase to approximately 500 kΩ, causing a 5-volt drop across R_E. With this large a drop, 6-volt battery operation would not be feasible.

14.6.4. Characterization of Transistor Noise. Although shot and thermal noise are predictable from noise theory with a high degree of accuracy, $1/f$ noise is not. Noise figure expressions should therefore be determined by experimental means. The noise specified is sometimes measured at a center frequency of 1000 cps, with an effective noise bandwidth of 1 cps. For most infrared preamplifier applications, where performance below 1000 cps is important, such a specification is useless. Much more important is the broadband noise figure, which can be considered a noise figure averaged over the passband of the amplifier.

The variation of the noise figure with source impedance is also important. In Fig. 14-14, an ideal noiseless amplifier, with equivalent noise current and noise voltage sources connected to the input, is substituted for a noisy amplifier. The equivalent rms input noise voltage e_n is found by measuring the output noise voltage with the

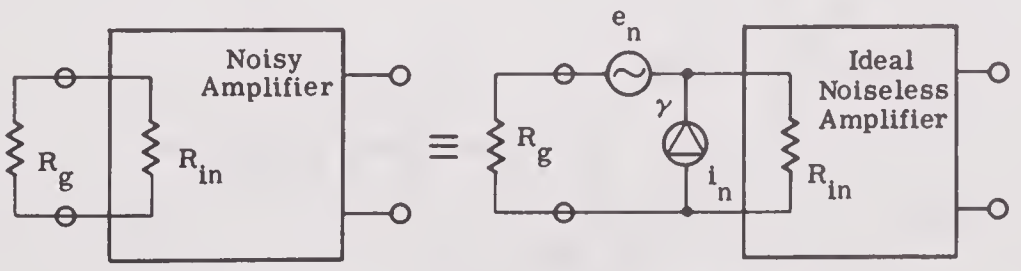


FIG. 14-14. Transistor equivalent noise generator.

input shorted, then dividing by the amplifier voltage gain. The equivalent rms input noise current i_n is found by measuring the output noise voltage with the input open, then dividing by amplifier voltage gain and input impedance. Since e_n and i_n are randomly fluctuating quantities, the degree of correlation between them must be known before the total power resulting from their combined effect can be computed. The degree of correlation is represented by a correlation coefficient, γ.

The noise factor is [13]

$$F = 1 + (4kTB)^{-1} [(e_n^2/R_g) + i_n^2 R_g + 2\gamma e_n i_n] \tag{14-91}$$

The correlation coefficient can range between 0 and 1. Neglecting the $1/f$ region, $\gamma = (h_{FE})^{-1/2}$ for low emitter currents. Thus, for large values of current gain, γ can be quite small. In the $1/f$ region, the correlation coefficient increases slightly.

If Eq. (14-91) is differentiated with respect to R_g and set equal to zero, the optimum value of R_g is found to be

$$R_{opt} = e_n/i_n \tag{14-92}$$

Substituting R_{opt} in Eq. (14-91), the minimum noise factor can be found to be

$$F_{min} = 1 + (1 + \gamma)\, e_n i_n / 2kTB \tag{14-93}$$

The values e_n and i_n are functions of I_E. Therefore, F is valid only at the bias condition at which e_n and i_n are measured. These two noise generators are fairly independent of collector voltage for voltages below 6 to 10 volts.

Assuming the noise sources are uncorrelated ($\gamma \simeq 0$),

$$F_{min} = 1 + e_n i_n / 2kTB \tag{14-94}$$

By combining equations, F can be expressed in terms F_{min} and R_{opt} as

$$F = 1 + (F_{min} - 1)\, K \tag{14-95}$$

where

$$K = \frac{1}{2}\left[\frac{R_g}{R_{opt}} + \frac{R_{opt}}{R_g}\right] \tag{14-96}$$

The value of K may be found in Fig. 14-15. As an example, assume the following conditions:

minimum noise figure = 1.5 db ($F_{min} = 1.4$)

$R_{opt} = 15$ kΩ

$R_g = 120$ kΩ

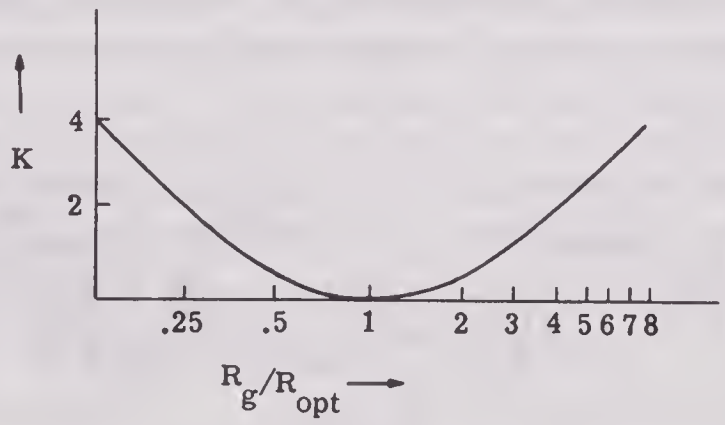


FIG. 14-15. Noise-figure bandwidth.

Since $R_g/R_{opt} = 8$, k is found from Fig. 14-15 to be 4. Therefore, the noise factor is

$$F = 1 + (1.4 - 1)(4) = 2.6 \tag{14-97}$$

and the noise figure NF is

$$NF = 10 \log_{10} 2.6 = 4.1 \text{ db} \tag{14-98}$$

Minimum noise figure and optimum generator resistance depend upon the frequency band to be utilized. Table 14-2 lists data for a typical 2N2176 low-noise silicon alloy transistor. Figure 14-16 shows the noise figure of the 2N2176 transistor as a function of source resistance for the three amplifier types. Figure 14-17 illustrates e_n and i_n as functions of emitter current [14].

TABLE 14-2. 2N2176 LOW-NOISE SILICON ALLOY TRANSISTOR DATA

Amplifier Passband	B	e_n	i_n	R_{opt}	F_{min}*	NF_{min}
0.8 - 50 cps	80 cps	0.16 μv	40 $\mu\mu$a	4k	3.0	4.75 db
f_0 = 1 kc	100 cps	0.056 μv	4.8 $\mu\mu$a	12k	1.68	2.25 db
0.8 - 10 kc	15 kc	0.7 μv	100 $\mu\mu$a	7k	2.17	3.3 db

$$I_c = 20 \ \mu\text{a}$$

$$V_{ce} = -1.5 \text{ v}$$

*Assuming noise generators have been fully correlated ($\gamma = 1$) for worst-case analysis.

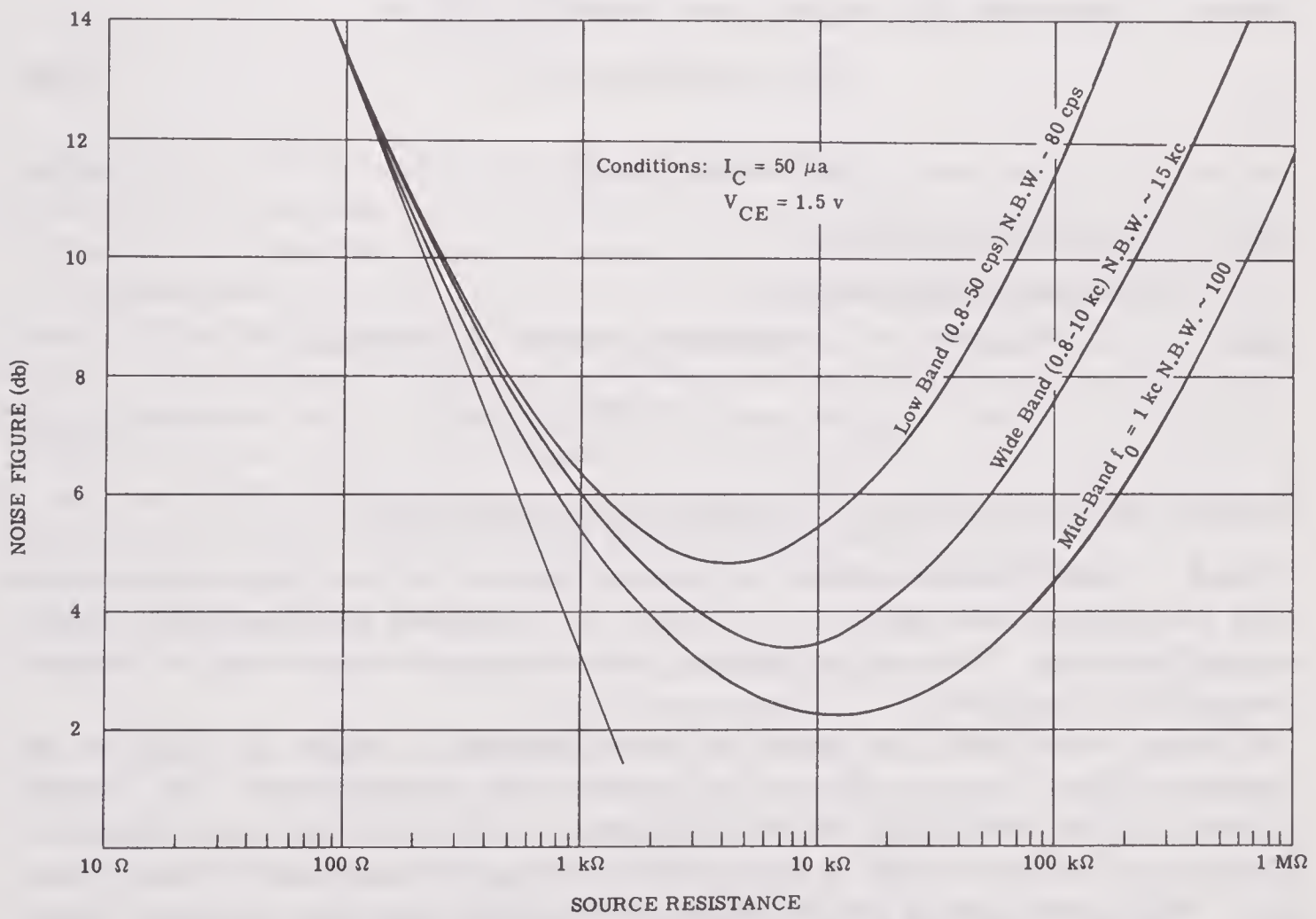


FIG. 14-16. 2N2176 transistor noise figure *vs.* source resistance.

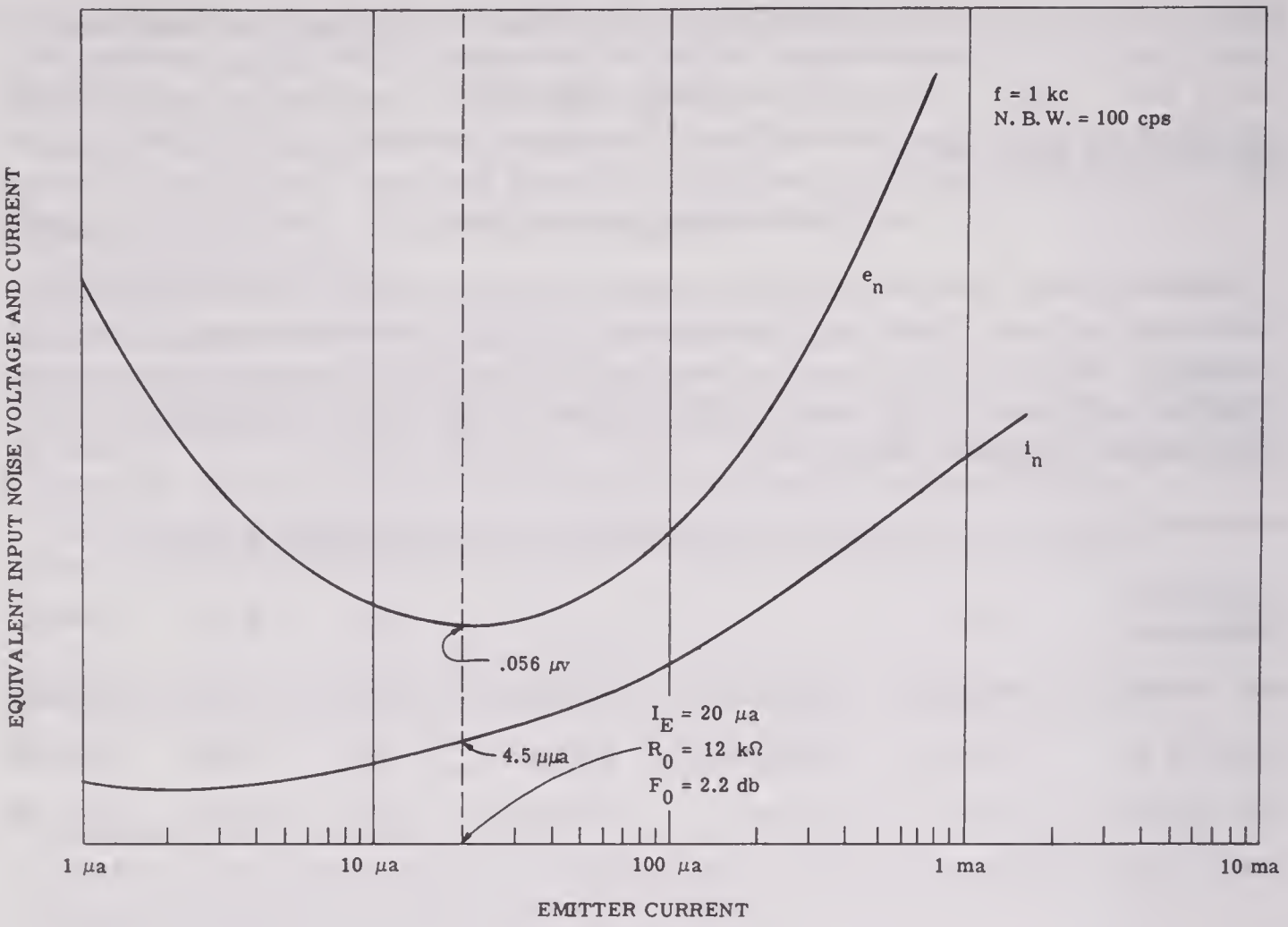


FIG. 14-17. 2N2176 transistor emitter current *vs.* equivalent input noise voltage and current.

14.6.5. Relationship of Minimum Noise Factor to Transistor Parameters. The minimum noise factor F_{min} can be related to transistor parameters in the following manner. In the white noise region, it can be shown [15] that

$$\overline{e_n^2} \approx 4kTB(r_e/2 + r'_b) \tag{14-99}$$

and

$$\overline{i_n^2} \approx 2e(I_c/h_{FE})B \tag{14-100}$$

where r_e = ac emitter resistance

r'_b = base spreading resistance

Since $r_e = kT/qI_E$ and $I_E \approx I_c$, substitution into Eq. (14-94) yields

$$F_{min} \approx 1 + \left[\frac{1 + (2e/kT)I_c r'_b}{h_{FE}}\right]^{1/2} \tag{14-101}$$

Therefore, as h_{FE} decreases or I_c increases, the noise increases.

14.6.6. Circuit Considerations. In infrared systems having a high signal-to-noise ratio, preamplifier noise factors of 10 to 20 db are acceptable, and preamplifier design is straightforward. However, in systems with low signal-to-noise ratios, a low-noise preamplifier is mandatory.

Minimum noise factor and optimum source impedance remain the same for the common emitter, common collector, or common base configurations. The common collector is often used in high-impedance designs because of its high input impedance. However, the common emitter is most desirable because of its greater available power gain. High power gain in the first stage is essential if noise from successive stages is to be minimized. The following factors should be considered in low-noise designs:

(1) A transistor specifically designed for low-noise and low-leakage current should be used.

(2) Because $1/f$ semiconductor noise is a major contributor to overall noise, the amplifier's low-frequency half-power point should be no lower than required by system considerations.

(3) The noise figure depends largely on emitter current (essentially collector current) and source impedance. In general, bias current for optimum operation occurs between 10 and 300 μa. However, plotting noise figure *vs* source impedance for a specific value of emitter current results in a rather broad minimum. Table 14-3 indicates the noise figure minima for modern low-noise planar transistors. For a source impedance far below 500 ohms, the best minimum noise figure is obtained by using a step-up transformer at the input stage. This method also provides additional voltage gain. Very low collector current is required if the minimum noise figure is obtained using a source impedance in the megohm region. Bias stabilization becomes difficult, so the transistor must have low collector leakage current if operation over a wide temperature range is required.

TABLE 14-3. NOISE FIGURE MINIMA FOR PLANAR TRANSISTORS

R_g (ohm)	I_E (μa)
0.5 – 3k	300
1.5 – 8k	100
3k – 15k	30
8k – 40k	10
15k – 200k	3
100k – 2m	1

(4) The fact that the α cutoff frequency of a transistor decreases with emitter current should also be considered. The narrowband noise figure starts rising at $f = f_{\alpha b}(1-\alpha_0)^{1/2}$. Therefore, it is important to use a transistor with a cutoff frequency much higher than the frequency desired. At low emitter (or collector) current, requirements for low-noise figure and the desired frequency response can necessitate the use of a higher bias current value. As an example, Table 14-4 lists data for the 2N2645 low-noise, high-gain, planar transistor having a 10-volt collector-to-base voltage. From the relationship $1-\alpha_0 \approx 1/h_{FE}$, the noise figure is found to rise at 13.5 kc for $I_c = 1$ μa; at 95 kc for $I_c = 10$ μa; and at 560 kc for $I_c = 100$ μa. Therefore, in the design of a 50-kc low-noise amplifier, the first stage would not be operated at a bias current of 1 μa.

TABLE 14-4. DATA FOR 2N2645 PLANAR TRANSISTOR

	I_c		
	1 μa	10 μa	100 μa
h_{FE}:	35	55	80
$f_{\alpha b}$:	80 kc	700 kc	5 Mc

(5) The bias point should be stabilized against temperature variations (Sec. 14.6.3). A stability factor S of 5 or less is usually necessary.

(6) The optimum noise figure of the transistor is always increased by the effect of the base divider network and any unbypassed emitter resistance. This effect is negligible if the following conditions are satisfied:

$$R_B \geq 10\, R_g \qquad R_E \leq 0.1\, R_g \tag{14-102}$$

where R_E = the external unbypassed emitter resistance. The minimum noise figure and the optimum source impedance of a transistor are essentially independent of any applied feedback. They are affected only by the circuit components which must be added to apply the feedback (such as an unbypassed emitter resistor) and their presence can only increase the noise figure. The effect of associated circuitry on noise performance is discussed in [16].

(7) Care should be taken when using Zener diodes in dc coupling and low-level biasing. Zener diodes generate large amounts of noise which can be reduced by bypassing the diode with a large capacitor; however, the Zener diode is not recommended for use in a low-noise-input stage which must respond to low frequencies.

14.6.7. Transformer Coupling. In many cases of preamplifier design, transformers provide an excellent solution to coupling problems presented by low- and high-impedance sources. For low-impedance sources, a step-up transformer can reduce amplifier stages and complexity. For high-impedance sources, a step-down transformer can be useful, provided the ratio between required high- and low-frequency response is no more than a few decades. Broadband transformer design becomes increasingly difficult at higher impedance levels. The lower the frequency to be passed, the larger the transformer must become to provide (a) the required magnetizing inductance, and (b) the necessary dynamic range. Enough iron must be used to maintain B_{max} well below saturation level for the largest input signal. The limiting factor is usually the magnetizing inductance required.

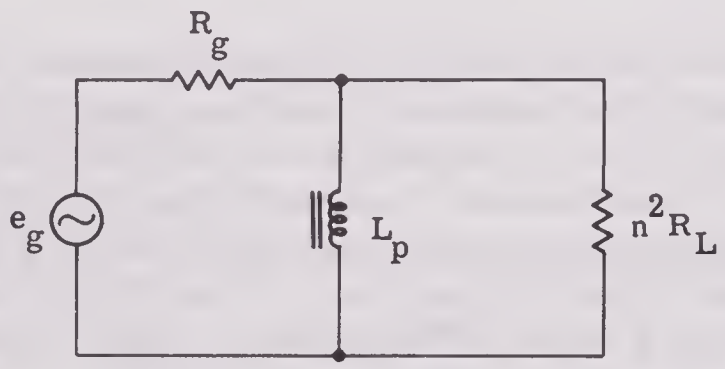


FIG. 14-18. Low-frequency equivalent circuit of a transformer.

The low-frequency equivalent circuit of a transformer is shown in Fig. 14-18. It is assumed that

$$r_p < R_G \qquad r_s < R_L$$

and the low-frequency 3-db point is given by

$$f_1 = \frac{n^2 R_L R_G}{2\pi L_p (R_G + n^2 R_L)} \tag{14-103}$$

Assuming that a close-to-optimum noise figure can be obtained by impedance matching a 1-MΩ detector to a preamplifier having a 10 kΩ input impedance, and that a one-cycle low-frequency response is desired, then

$$n^2 R_L = R_G = 1 \text{ M}\Omega$$

$$f_1 = 1 \text{ cps}$$

Solving for the required primary inductance

$$L_p = \frac{10^6 \times 10^6}{2\pi(1)(2 \times 10^6)} = 8000 \text{ hy} \tag{14-104}$$

Such a value is unreasonable. However, for a 1000-cps response, a realistic 8 henries is required.

Transformers usually have appreciable stray capacitance because of the great number of windings over a small area. This limits the high-frequency response. Input transformers must operate at very low power levels; consequently, they should be enclosed in a magentic shield, especially if the preamplifier is located close to a chopper motor or power transformer. The transformer should also contain an electrostatic (Faraday) shield between primary and secondary windings to minimize stray capacitance coupling between windings. The inclusion of such a shield, however, increases the transformer shunt capacitance, which, again, lowers the high-frequency response. Transformer windings should be mechanically rigid to prevent microphonics caused by minute capacitance changes. Careful encapsulation will generally reduce transformer microphonics to a negligible level.

14.6.8. Simple Low-Noise Preamplifier Design [17]. The amplifier shown in Fig. 14-19 has a voltage gain of approximately 20 db with a flat response from 10 cps to 50 kc. By referring to Table 14-3, it can be seen that for $R_g = 300$ kΩ, and a bias current of about 1 μa, a low-noise figure is obtained. The collector current of the input transistor is 1.5 μa. The common emitter stage is biased for optimum noise performance and provides good power gain. The input impedance is 1.2 MΩ up to

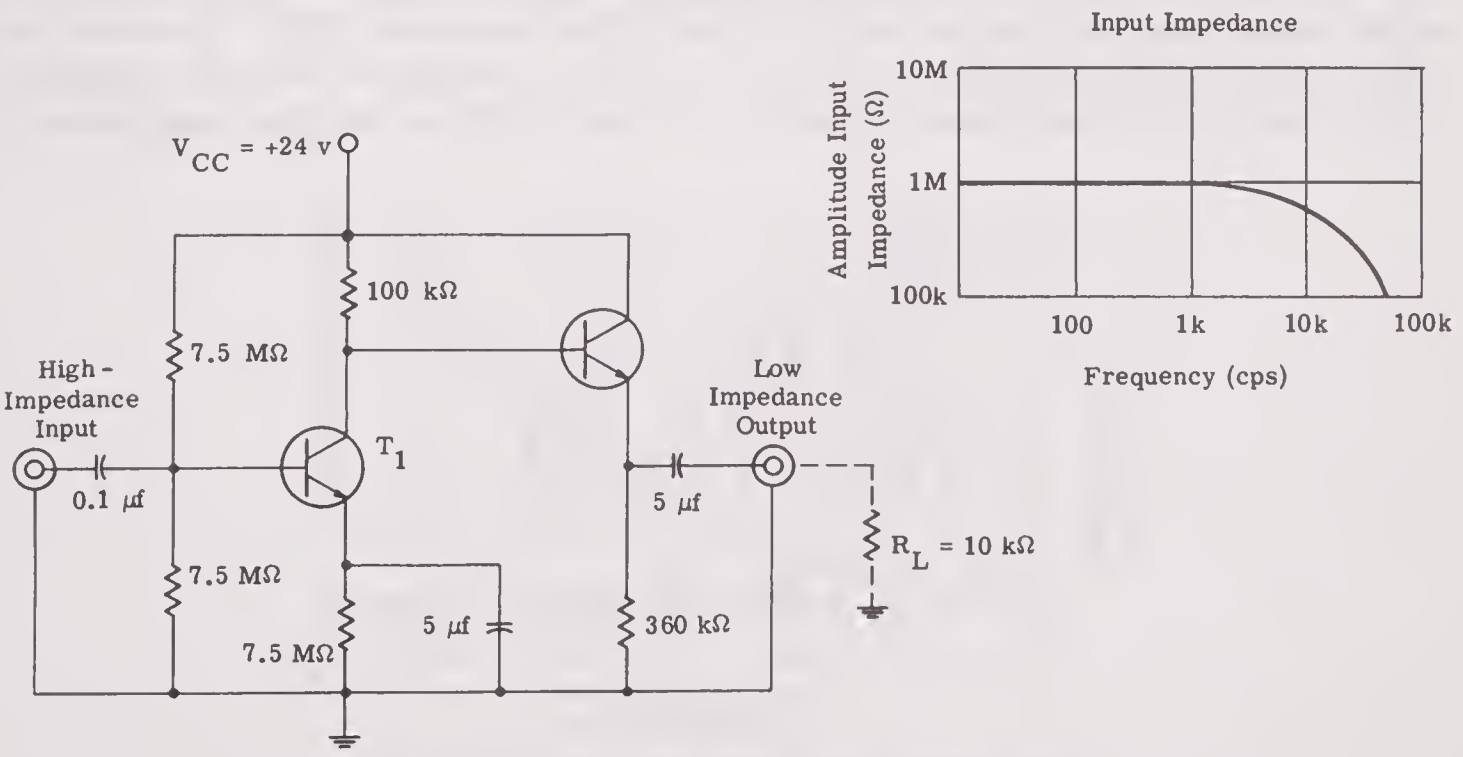


FIG. 14-19. High-impedance wideband amplifier with 2N2484 transistors.

4 kc, and the narrowband noise figure is 1.6 db at 1 kc. The stability factor is approximately 1.5 and is given by

$$S \simeq 1 + R_B/R_E \tag{14-105}$$

where $R_B = 3.75$ MΩ

$R_E = 7.5$ MΩ

The characteristics of typical low-noise transistors are listed in Table 14-5.

TABLE 14-5. CHARACTERISTICS OF TYPICAL LOW-NOISE TRANSISTORS

Type	*Wideband* Noise Figure* $\overline{NF}$	*Collector Leakage Current* I_{CBO}	*dc Current Gain,* h_{FE}	*Additional Data*
2N2484	1.8 db typ 3 db max	10 na max at $V_{CB} = 45$v	30 min at $I_c = 1$ μa 100 min at $I_c = 10$ μa 200 min at $I_c = 500$ μa	$h_{fe} = 15$ min at $F = 1$ mc, for $I_c = 50$ μa
2N2645	3.5 db max	0.4 na typ 10 na max at $V_{CB} = 60$v	20 min at $I_c = 10$ μa 60 min at $I_c = 100$ μa	$h_{fe} = 2.5$ min at F = 20 mc, for $I_c = 10$ ma $f_{ab} = 0.7$ mc (typical) for $I = 10$ μa
2N930	4 db max	2 na max	100 min at $I_c = 10$ μa 150 min at $I_c = 500$ μa	$h_{fe} = 1$ min at F = 30 mc, for $I_c = 500$ μa
2N2586	1.5 db typ†	2 na max at $V_{CB} = 45$v	80 min at $I_c = 1$ μa 120 min at $I_c = 10$ μa 150 min at $I_c = 500$ μa	$h_{fe} = 1.5$ min at F = 30 mc, for $I_c = 500$ μa
2N2524	2 db typ 3 db max	2 na max 1 na typ at $V_{CB} = 45$v	60 min at $I_c = 1$ μa 100 min at $I_c = 10$ μa 150 min at $I_c = 500$ μa	$h_{fe} = 1.5$ min at $F = 30$ mc, for $I_c = 500$ μa

*Power bandwidth of 15.7 kc, 3-db points at 10 cps and 10 kc, $V_{CB} = 5$ v, $I_c = 10$ μa, $R_g = 10$ k.
†Narrowband data indicate the same $\overline{NF}$ is obtained for $I_c = 1$ μa, $R_g = 1$ Meg.

14.7. Field-Effect Transistors

14.7.1. Noise Figure. Although good noise figures can be obtained by using conventional planar transistors with source impedances up to 1 MΩ, at higher impedances better results can be obtained with the field-effect transistor (FET). In controlled FET's, the $1/f$ corner frequency is less than 100 cps, almost an order of magnitude lower than most conventional transistors. Figure 14-20 shows the noise figure of

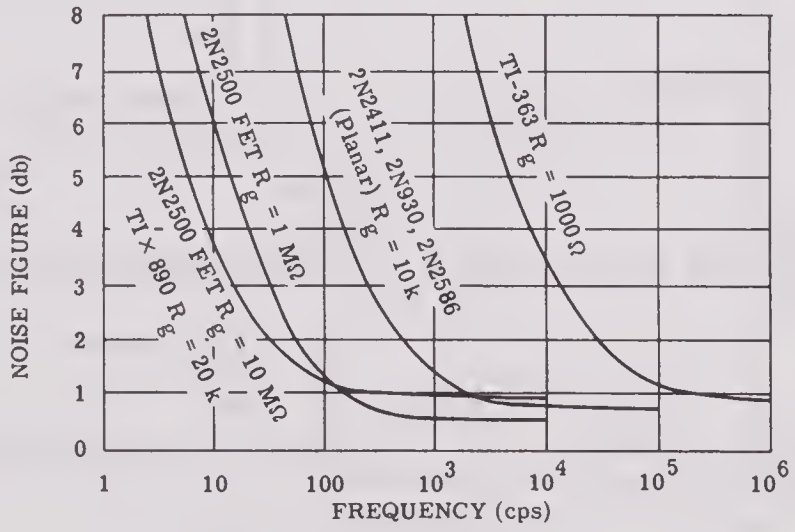


FIG. 14-20. Field effect transistor noise figure *vs.* frequency [18].

several FET's as a function of frequency [18]. The FET provides lower noise figures in the low-frequency region, and permits low-noise designs for source impedances well into the megohm region.

Figure 14-14 shows how FET noise can be characterized by an equivalent noise voltage generator and an equivalent noise current generator. In the case of the FET, i_n is very small, leading to a high optimum source impedance. The narrowband values of these equivalent generators are shown in Fig. 14-21 as a function of frequency, for a low-noise 2N2500 FET. The variation of optimum R_g and optimum *NF vs.* frequency appears in Fig. 14-22. Stability can be obtained in FET's by using self bias.

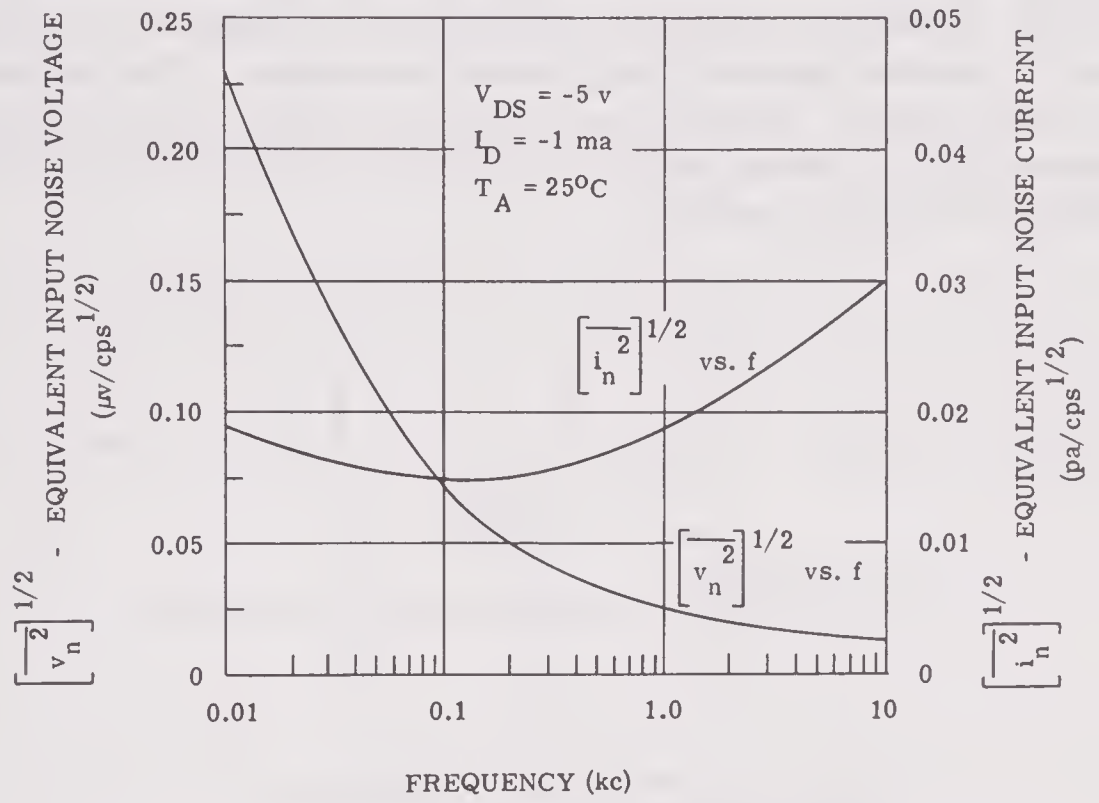


FIG. 14-21. Equivalent input noise voltage and equivalent input noise current *vs.* frequency [18].

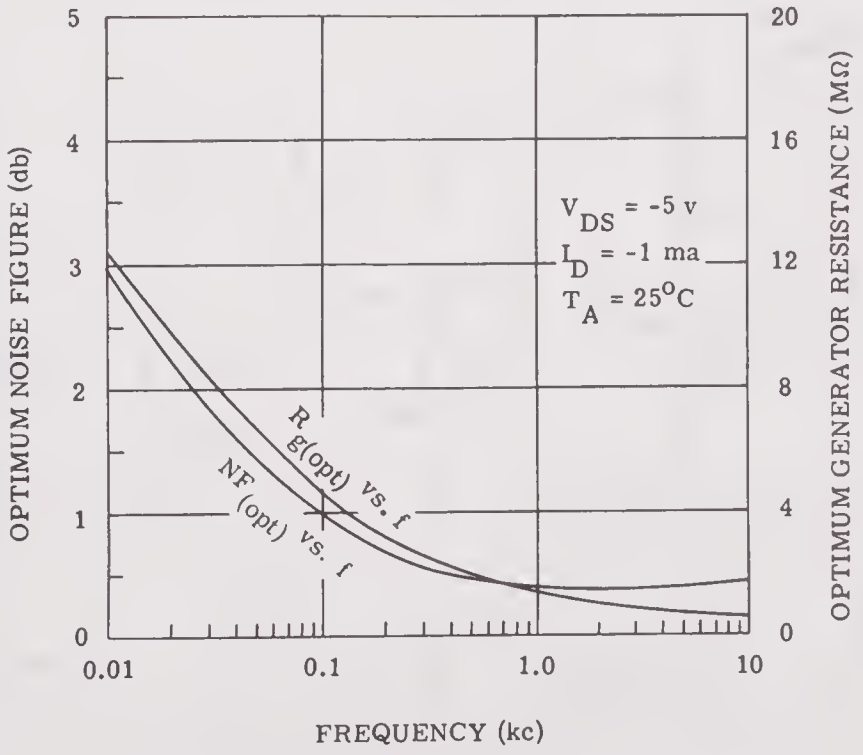


FIG. 14-22. Optimum noise figure and optimum generator resistance *vs.* frequency [19].

14.7.2. FET Preamplifier Design. Figure 14-23 shows that excellent results for high-impedance preamplifiers can be obtained by combining the FET with a planar transistor in a cascode configuration. The Miller effect is reduced in this arrangement, thereby increasing the bandwidth of a low-noise preamplifier. In addition, independent adjustment of the transistor's operating conditions permits optimum noise performance.

Figure 14-24 shows the application of two FET's in a low-noise, high-impedance amplifier [19]. The frequency response of the amplifier is shown in Fig. 14-25, and the broadband noise figure *vs.* source impedance is given in Fig. 14-26. The amplifier has a fixed gain of 100 and an input impedance of 30 MΩ shunted by 8 pf. The broadband noise factor is less than 3 db with a generator resistance of 50 kΩ to 5 MΩ. The 2N2498 FET is operated at a drain current of 1 ma, and the 2N930 transistor is operated at $I_c = 100$ μa. Each bias current is optimum for the respective transistor. The second stage is operated common base (cascode connection). The optimum source impedance is the same for common base or common emitter. The 10 kΩ optimum source impedance for the second stage is provided by paralleling the 20 kΩ emitter resistor with the 20 kΩ drain resistor.

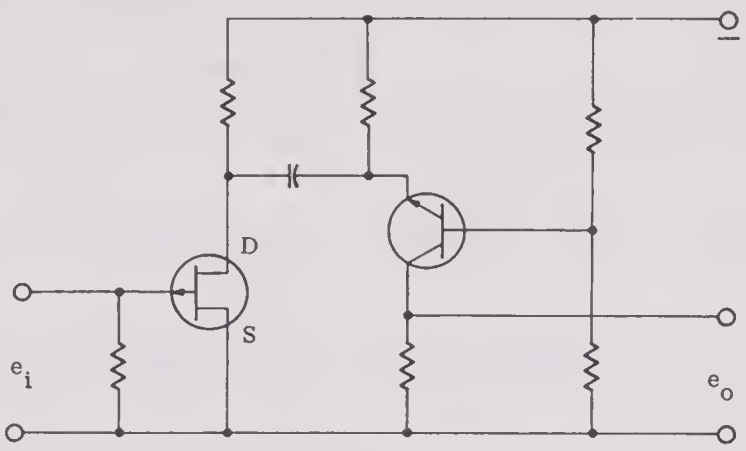


FIG. 14-23. Field-effect transistor and planar transistor in cascode.

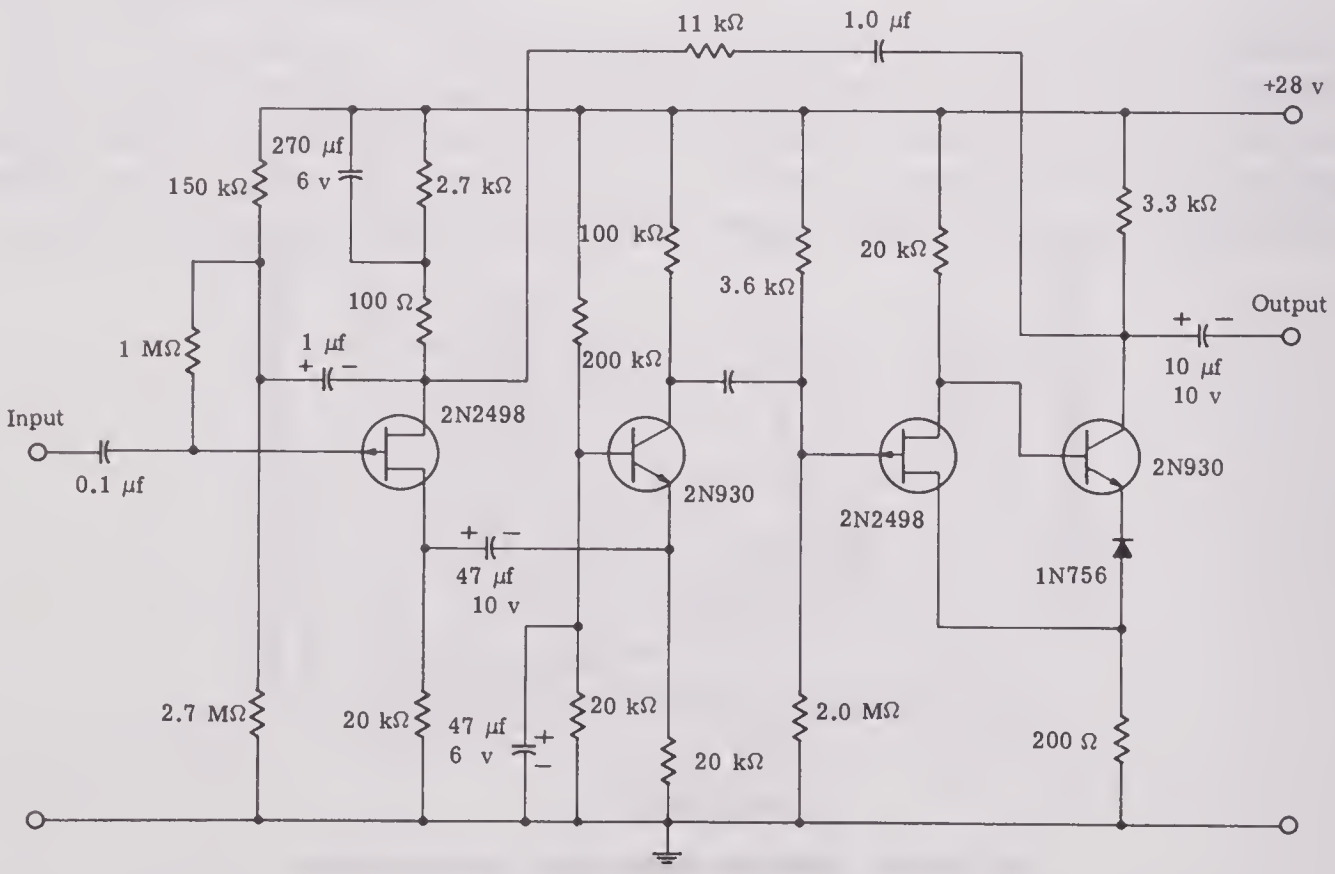


FIG. 14-24. FET low-noise, high-impedance amplifier [19].

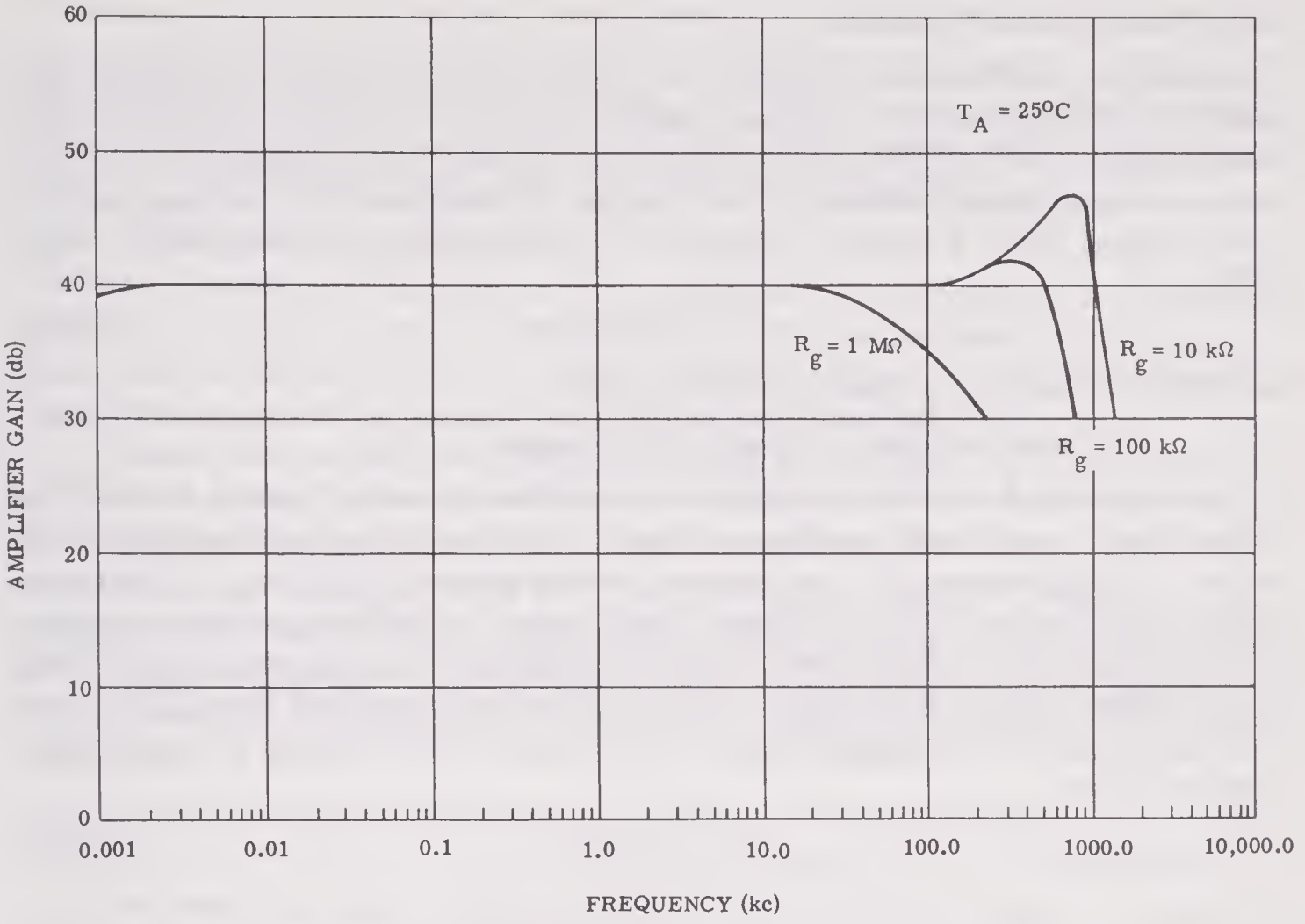


FIG. 14-25. Frequency response of FET amplifier [19].

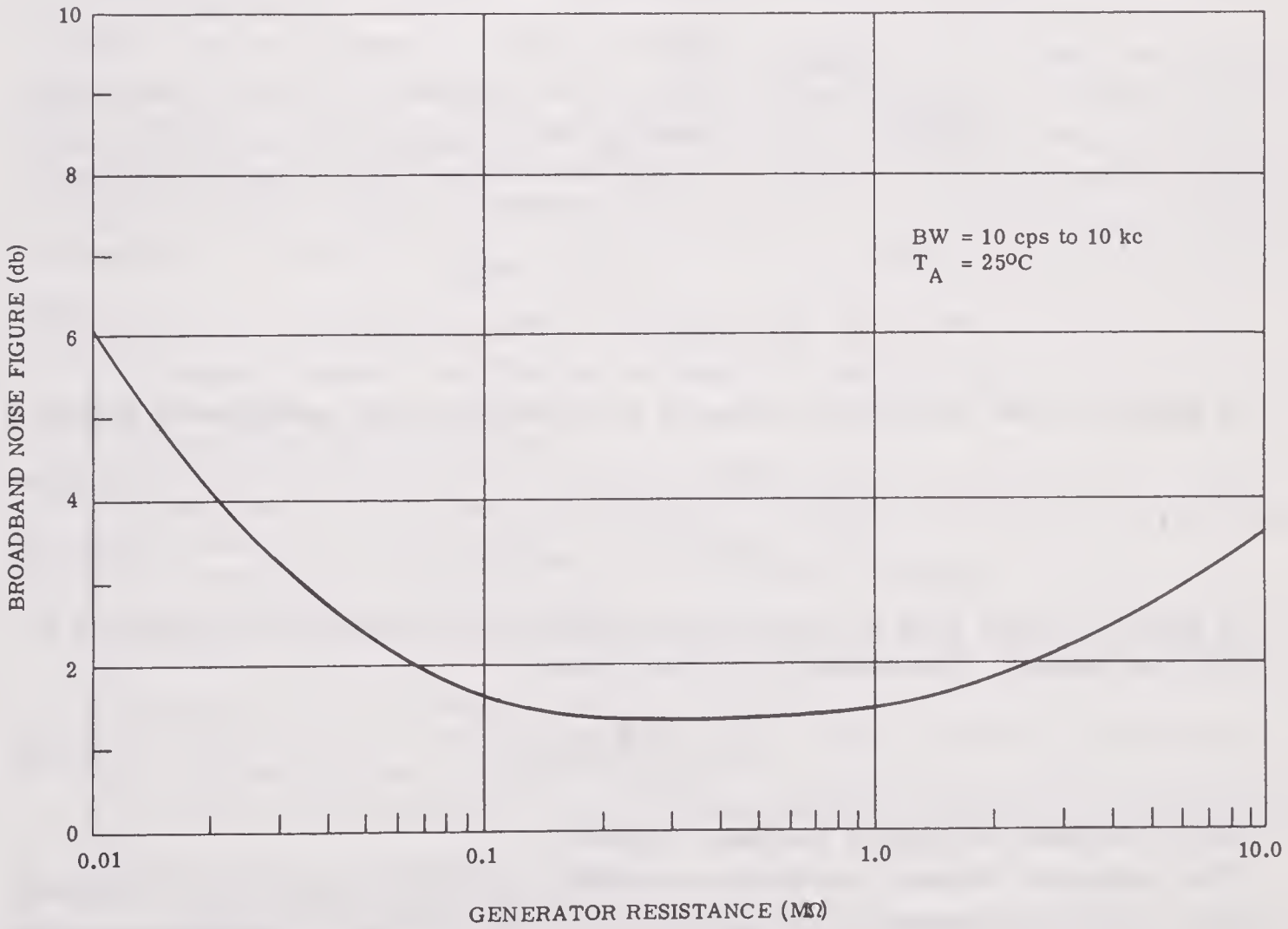


FIG. 14-26. FET broadband noise figure *vs.* source impedance [19].

14.8. Synchronous Detection

Radiometric instruments frequently use choppers to interrupt the radiation periodically in the optical path. Choppers permit operation in the high-frequency region where detector and amplifier characteristics are better. The frequency of the amplitude-modulated signal produced by the chopper is determined by the chopping rate. If the chopper blade is shaped to produce sine wave modulation, the detector output will be

$$e_s = \tfrac{1}{2} E_s(t) \sin \omega t \tag{14-106}$$

where $E_s(t)$ = the peak-to-peak signal amplitude

ω = angular frequency produced by chopper

A synchronous detector is essentially a narrowband detection system in which the target signal is beat with a reference signal of the same frequency producing a dc output. A block diagram of a synchronous detector is shown in Fig. 14-27. The output of the target detector is an intelligence signal which is amplified and then multiplied by a reference signal. The reference signal is generated by a magentic or photoelectric device synchronized by the chopper. Therefore, the frequency and phase of the reference signal and the intelligence signal are the same. The output of the reference detector is then

$$e_r = E_r \sin \omega t \tag{14-107}$$

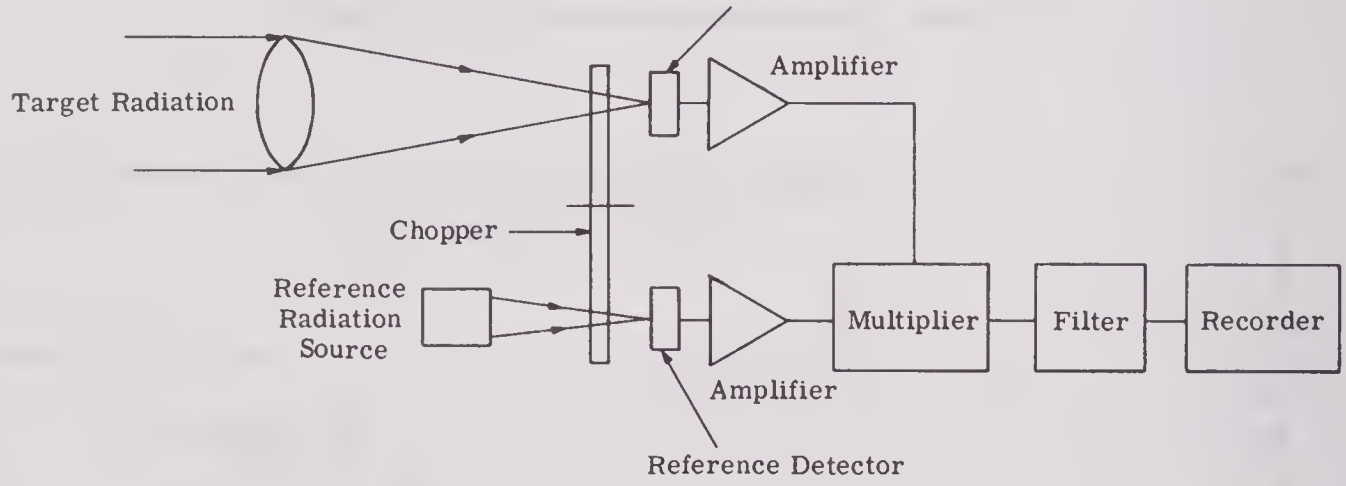


FIG. 14-27. Block diagram of synchronous detector.

The product of the two signals appearing at the output of the multiplier is therefore

$$e_m(t) = KE_s(t)E_r \sin^2 \omega t \tag{14-108}$$

or

$$e_m(t) = KE_s(t)E_r (\cos 2\omega t - 1) \tag{14-109}$$

The passband of the filter is made much less than 2ω to eliminate the unwanted frequency component. The output of the filter is then

$$e_0(t) = \frac{KE_r}{2} E_s(t) \tag{14-110}$$

Thus the original intelligence has been recovered.

The similarity between synchronous detection and cross correlation is discussed in [20]. The improvement in the signal-to-noise ratio obtained by synchronous rectification depends upon (a) the form and frequency of the target signal, (b) the form of

the reference signal, and (c) the power spectrum of the noise entering the correlator. If one assumes that the radiation and the reference signals vary sinusoidally with time and that the system suffers from band-limited white noise, a figure of merit $Q(T)$ can be derived by determining the ratio of the signal-to-noise ratio at the output of the filter to the signal-to-noise ratio at the input to the preamplifier.

14.9. Grounding Considerations

14.9.1. Grounding Techniques. The grounding of preamplifiers and associated circuits is of major importance. Improper grounding can cause self-sustained oscillations, gain distortions, and numerous other undesirable effects. The following grounding techniques should be followed in all design considerations:

(1) Ground the amplifier circuit to the chassis at the point of lowest signal level.

(2) Avoid ground loops. Avoid grounding shielded leads at both ends. Tie the power supply return to the amplifier at the point of lowest signal level.

(3) Do not use a single ground bus for multiple-stage amplifiers; instead, provide a ground bus for every amplifier stage having a gain of 100 or more. Return grounds of all stages to the point of lowest signal level. If the ground for the highest level stages cannot be returned to this point, then a separate power supply return should be provided.

(4) Use large diameter wire for ground returns. For low-level stages, the minimum size should be No. 22 or No. 20; for higher level stages, No. 18 is the minimum size.

(5) Separate the ground leads of low-level and high-level stages. If necessary, shield the high level ground lead, then ground the shield at the point of lowest signal level.

(6) Keep currents out of all shields by tying the shield to the ground at the point of lowest signal level.

14.9.2. Low-Noise Cable. A major problem in the design of low-noise preamplifiers is the spurious audio frequency noise generated in coaxial cables due to shock, excitation, or vibration. This cable noise can completely mask the desired signal, particularly in high-impedance circuits subjected to shock or vibration. The mechanism of noise in coaxial cables is discussed in [21]. Miniature low-noise cable is now available commercially from Microdot, Inc., under the trade name of Mininoise Coax Cable [21].

References

1. W. R. Bennett, *Electrical Noise,* McGraw-Hill Book Co., New York, 1960.
2. J. B. Johnson, *Phys. Rev.* **32,** 97 (1928).
3. H. Nyquist, *Phys. Rev.* **33,** 110 (1928).
4. W. Schottky, *Ann. Phys.* **57,** 541 (1918).
5. W. B. Davenport and W. R. Root, *An Introduction to the Theory of Random Signals and Noise,* McGraw-Hill Co., New York, 1958.
6. B. J. Thompson, D. O. North, and W. A. Harris, *RCA Rev.* **1,** 269 (1940).
7. K. M. van Vliet, *Proc. IRE* **46,** 1004 (1958).
8. S. Goldman, *Frequency Analysis Modulation and Noise,* McGraw-Hill Book Co., New York, 1948.
9. A. van der Zeil, *Proc. Nat. Electron. Conf.* **17,** 454 (1961).
10. G. E. Valley, Jr., and H. Wallman, *Vacuum Tube Amplifiers,* Radiation Laboratory Series Vol. 18, McGraw-Hill Book Co., New York, 1948.
11. G. M. Glasford, *Fundamentals of Television Engineering,* McGraw-Hill Book Co., New York, 1955.
12. H. Wallman *et al., Proc. IRE* **36,** 700 (1948).
13. H. F. Cook, *Proc. IRE* **50,** 2520 (1962).
14. *The Evaluation of Transistor Low-Frequency Noise Characteristics,* Sperry Semiconductor Technical Application Bulletin No. 2110, Sperry Gyroscope Co., Great Neck, L.I., New York.

15. N. H. Martens, *Solid State Design,* **3,** 35 (1962).
16. R. D. Middlebrook and C. A. Mead, *Semiconductor Products,* **2,** 26 (1959).
17. *Fairchild Semiconductor Application,* Bull. No. 112, Fairchild Semiconductor Corp., Mountain View, Calif.
18. *Low Noise Seminar Papers,* Texas Instruments Incorporated, Dallas, Texas, 1963.
19. *Field Effect Transistor Theory and Application,* Texas Instruments Incorporated, Dallas, Texas.
20. T. J. Wieting, "Correlation Techniques for Infrared Detection Systems," *Proc. IRIS,* **V,** 3, 57 (1960).
21. *Coaxial Cable Review,* Microdot, Inc., 220 Pasadena Ave., South Pasadena, Calif.

Chapter 15

OPTICAL FREQUENCY-RESPONSE TECHNIQUES

R. Barakat

Itek Corporation

CONTENTS

15.1. Optical Systems and Linear-System Theory 614
15.1.1. General Concepts of Linear-System Theory 614
15.1.2. Necessary Formulas of Optical Diffraction Theory 615
15.1.3. Transfer Functions of Optical Systems in Incoherent Light . . . 621
15.2. Resolution and Its Ramifications 627
15.2.1. Resolution Criteria for Point Sources 627
15.2.2. Sine-Wave Resolution 628
15.3. Effect of Aberrations on Transfer Function 629
15.3.1. Spherical Aberration and Coma. 629
15.3.2. Computation of Transfer Function of an Actual System. 634
15.4. Effect of Apodization on the Transfer Function. 636
15.4.1. Altering High- or Low-Frequency Response by Apodization . . . 636
15.4.2. Luneberg Apodization Problems 636
15.5. Merit Factors 637
15.5.1. Linfoot Quality Factors 637
15.5.2. Plane of Best Focus 638
15.6. Frequency Response Calculations from Lens Design Data. 640
15.6.1. Computer Calculations of Frequency Response. 640
15.7. Optical Response Measuring Equipment. 643

15. Optical Frequency-Response Techniques

15.1. Optical Systems and Linear-System Theory

In large measure the optical terminology of linear-system theory is a transcription of the corresponding electrical terminology with minor modifications. Even though optical systems are two dimensional and operate in the spatial domain in contrast to electrical systems which are generally one dimensional in the time domain, the basic analysis is not altered.

15.1.1. General Concepts of Linear-System Theory. In a linear system the relation between the input $y_i(t)$ and output $y_0(t)$ is

$$y_0 = \mathfrak{L}[y_i] \tag{15-1}$$

where $\mathfrak{L}$ is an operator characterizing the system and t represents spatial coordinates. A system is linear if the following conditions are satisfied:

(1) Commutative condition: if y_i and z_i are two inputs, then

$$\mathfrak{L}[y_i + z_i] = \mathfrak{L}[z_i + y_i] \tag{15-2}$$

(2) Superposition condition:

$$\mathfrak{L}[y_i + z_i] = \mathfrak{L}[y_i] + \mathfrak{L}[z_i] = y_0 + z_0 \tag{15-3}$$

(3) Proportionality condition: If α is a constant (possibly complex), then

$$[\alpha y_i] = \alpha\mathfrak{L}[y_i] = \alpha y_0 \tag{15-4}$$

If h is the impulse response of the system, the superposition integral is

$$y_0(t) = \int_{-\infty}^{\infty} h(t', t)y_i(t')\,dt' \tag{15-5}$$

If the system is stationary (time invariant), then

$$y_0(t + \tau) = \mathfrak{L}[y_i(t + \tau)] \tag{15-6}$$

In the spatial domain such a system is called spatially invariant or isoplanatic. When stationary conditions are imposed the superposition integral is

$$y_0(t) = \int_{-\infty}^{\infty} h(t - \tau)y_i(\tau)\,d\tau \tag{15-7}$$

Denoting by capital letters the Fourier transforms of Eq. (15-7),

$$Y_0(f) = H(f)Y_i(f) \tag{15-8}$$

where f is the frequency (either time or spatial) and $H(f)$ is the transfer function. The system can be specified by the impulse response function (in the time domain) or the transfer function in the frequency domain.

15.1.2. Necessary Formulas of Optical Diffraction Theory.

Let us define two functions:

$a(x, y)$ = complex amplitude of diffracted image

$t(x, y)$ = point spread function

where (x,y) = coordinates in receiving plane. $a(x,y)$ is not an observable quantity in the optical and infrared regions (at least by present techniques); however, its absolute square $t(x,y)$ is observable:

$$t(x, y) = |a(x, y)|^2 \tag{15-9}$$

or

$$i'(x, y) = Ct(x, y) \tag{15-10}$$

where $i'(x,y)$ is the illuminance (flux per unit area). Since $t(x,y)$ is dimensionless, the constant C has the necessary dimensions.

The normalized illuminance ratio is

$$i(x, y) = \frac{t(x, y)}{t_{A_i}(0, 0)} \tag{15-11}$$

where $t_{A_i}(0,0)$ denotes the illuminance at $x = y = 0$ (geometric center of diffraction pattern) for an aberration-free Airy objective with no losses. The $i(x,y)$ expresses the distribution of illuminance in the diffracted image as a function of that at the center of the aberration-free Airy aperture and is a ratio (no dimensions). In practice $i(x,y)$ is called the illuminance (or intensity).

15.1.2.1. *Kirchhoff Diffraction Theory.* Kirchhoff diffraction theory states that the complex amplitude due to a point source lying on the optical axis of the system is

$$a(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(\xi, \eta) \exp\left[i \frac{k}{f}(\xi x + \eta y) \right] d\xi \, d\eta \tag{15-12}$$

where $k = 2\pi/\lambda$ = wave number of incident light

f = focal length of the system

ξ, η = rectangular coordinates in exit pupil plane

The quantity $A(\xi,\eta)$ or its equivalent $A(\beta\gamma)$ is defined below. The infinite range of integration is for formal simplicity; actually the region of integration is the area of the aperture. In normalized coordinates (where X and Y are dimensionless in the receiving plane),

$$X = \frac{2\pi\mu_0}{\lambda} x; \qquad Y = \frac{2\pi\nu_0}{\lambda} y \tag{15-13}$$

where μ_0 and ν_0 are the aperture semi-angles in the respective directions. The dimensionless angular coordinates in the aperture are

$$\beta = \frac{\xi}{\mu_0 f} = \frac{\mu}{\mu_0}; \qquad \gamma = \frac{\eta}{\nu_0 f} = \frac{\nu}{\nu_0} \tag{15-14}$$

where μ and ν are angular coordinates and the distance f is large enough so that the tangent of an angle can be approximated by the angle. In dimensionless coordinates,

$$a(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(\beta, \gamma) \, e^{i(\beta X + \gamma Y)} \, d\beta \, d\gamma \tag{15-15}$$

15.1.2.2. *Pupil Function.* The quantity $A(\beta,\gamma)$ is termed the pupil (or transmission) function and

$$A(\beta,\gamma) = \frac{\text{amplitude distribution of wave over exit pupil}}{\text{amplitude distribution of wave over entrance pupil}}$$

In the general case the pupil function is complex. $A(\beta,\gamma) \leq 1$ for physically realizable systems and must vanish outside the aperture.

$$\begin{aligned} A(\beta,\gamma) &= A_0(\beta,\gamma)\, e^{i2\pi W(\beta,\gamma)} \qquad (\beta,\gamma) \text{ is in the aperture} \\ &= 0 \qquad\qquad\qquad\qquad\quad (\beta,\gamma) \text{ is not in the aperture)} \end{aligned} \tag{15-16}$$

where $A_0(\beta,\gamma)$ = the real amplitude distribution over exit pupil

$W(\beta,\gamma)$ = the wave-front aberration in wavelength units. The wave-front aberration is also called optical difference, aberration function, and eikonal.

Fourier transform theory applied to Eq. (15-15) yields

$$A(\beta,\gamma) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a(X,Y)\, e^{-i2\pi(\beta X+\gamma Y)}\, dX\, dY \tag{15-17}$$

This equation is integrated over the entire Fraunhofer receiving plane. Thus the complex amplitude $a(X,Y)$ in the Fraunhofer receiving plane is the Fourier transform of the pupil function $A(\beta,\gamma)$. A knowledge of either $a(X,Y)$ or $A(\beta,\gamma)$ is sufficient to compute the other, but a knowledge of only $t(X,Y)$ is not sufficient since phase information is lost. There are an infinite number of pupil functions having the same spread function. The evaluation of $a(X,Y)$ for an aperture of arbitrary shape is very involved; however, for circular and rectangular apertures expressions are available for $a(X,Y)$.

15.1.2.3. *Circular Aperture.* For a circular aperture

$$a(v) = \int_0^1 A(\rho) J_0(v\rho)\rho\, d\rho \tag{15-18}$$

where $\rho = r/r_m$

$r^2 = (\beta^2 + \gamma^2)f^2$

r_m = radius of aperture

$v = \dfrac{2\pi r_m}{\lambda} \sin\alpha \approx \dfrac{2\pi r_m}{\lambda} \cdot \dfrac{z}{f}$

z = distance in receiving plane from center

f = focal length

J_0 = Bessel function of order zero

In the special case of an aberration-free Airy objective with no losses, $A(\rho) = 1$ and straightforward integration yields

$$i(v) = [2J_1(v)/v]^2 \tag{15-19}$$

This is the familiar Airy pattern of physical optics (Fig. 15-1).

15.1.2.4. *Annular and Annulus Apertures.* The annular aperture is a circular aperture in which the center is blocked out; the annulus aperture has a ring-shaped obstruction (Fig. 15-2). The annular aperture is a special case of the annulus aperture where ϵ' is 0.

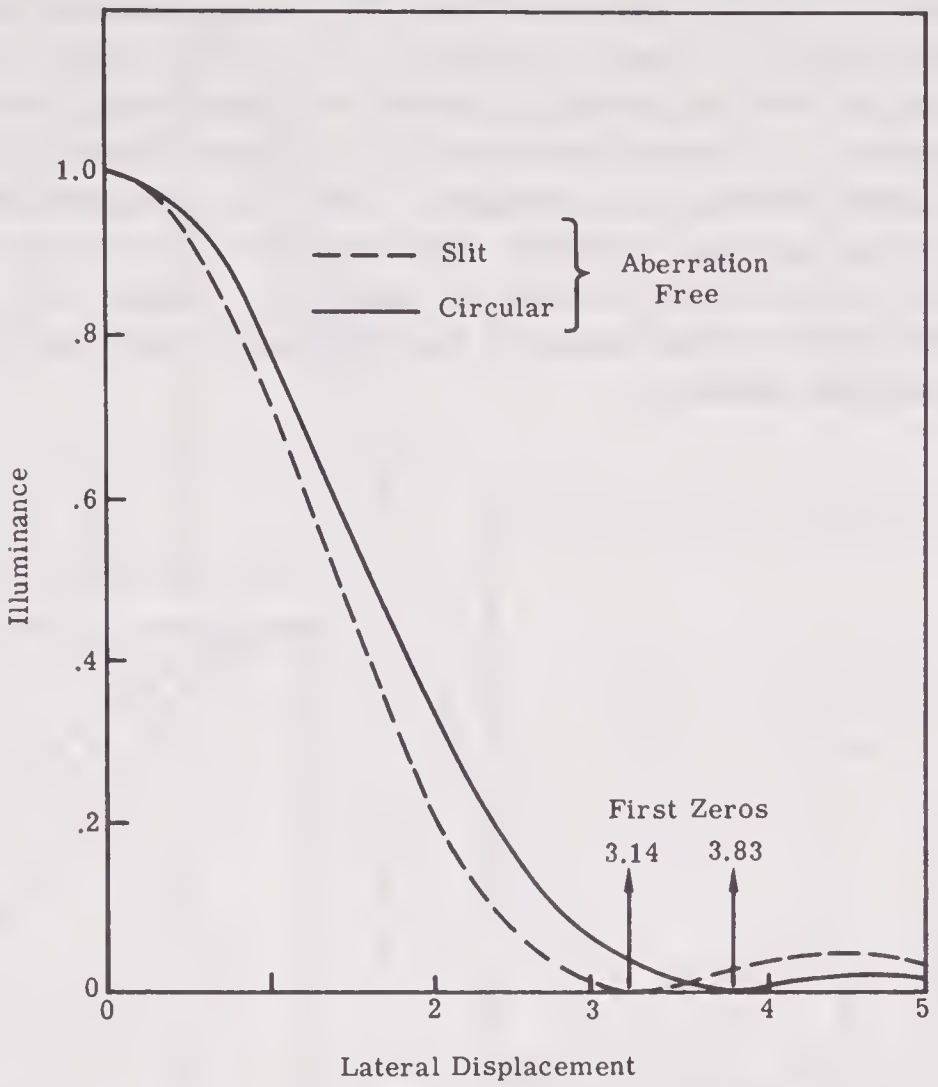
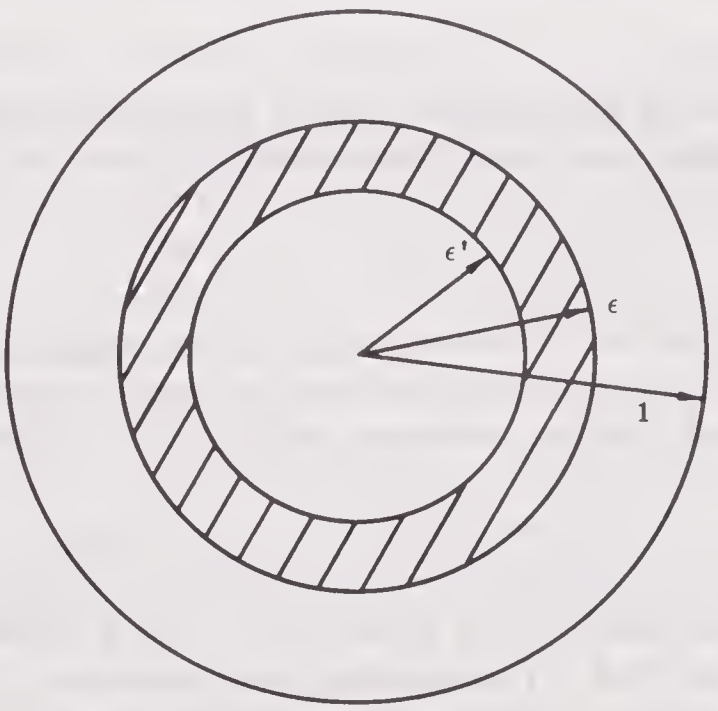


FIG. 15-1. Spread function for circular and slit apertures in Fraunhofer receiving plane.



$$0 \leq \epsilon' < \epsilon \leq 1$$

FIG. 15-2. Geometry of the annular and annulus apertures.

The effect of the annular aperture is to decrease the radius of the first minimum of the diffraction pattern from its value $v = 3.832$. The extent to which the minimum is decreased depends upon the amount of blocking of the central region of the aperture. The resultant gain in resolving power is obtained at a considerable loss of intensity in the diffraction pattern. A further constraint is the increased intensity of the secondary maxima as the central obstruction is increased. See [1] for a comprehensive treatment of the aberration-free annular aperture. Both annular and annulus apertures possessing spherical aberration are covered in detail [2]. Figure 15-3 illustrates the variation in distribution of illuminance in the Fraunhofer receiving plane for various aberration-free annular apertures.

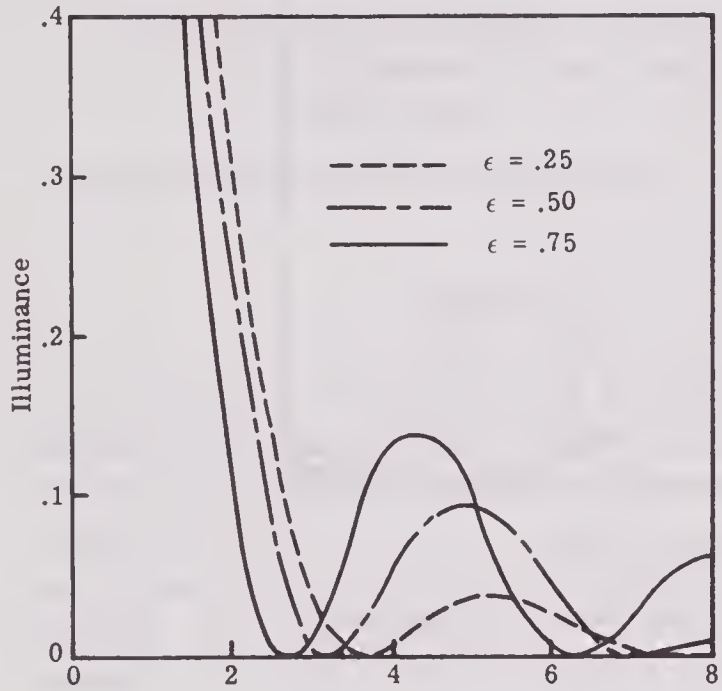


FIG. 15-3. Distribution of illuminance for aberration-free annular aperture in Fraunhofer receiving plane.

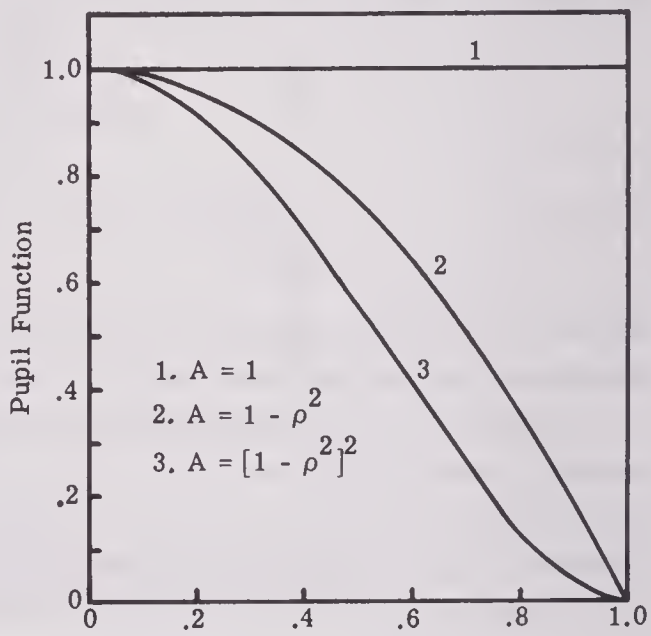


FIG. 15-4. Straubel pupil function for values of $n = 0, 1, 2$.

15.1.2.5. *Straubel Pupil Function for Circular Aperture.* A case of interest, because it admits of exact integration, is an aberration-free system with a pupil function of the form

$$A(\rho) = (1 - \rho^2)^n \qquad (n = 0, 1, 2, \ldots) \tag{15-20}$$

For $n = 1$ the pupil function follows a parabolic law roughly equivalent to a cosine distribution; for $n = 2$, the $A(\rho)$ is approximately a cosine squared distribution. The behavior of these pupil functions is shown in Fig. (15-4). The illuminance ratio is

$$i(v) = \left[2^{n+1} n! \frac{J_{n+1}(v)}{v^{n+1}} \right]^2 \qquad (n = 0,1,2,\ldots) \tag{15-21}$$

The Straubel curves include the Airy system ($A = 1$) as a subcase. The illuminance curves are shown in Fig. 15-5. To emphasize the differences in the patterns the maximum values of each pattern have been normalized to unity to permit comparison. Increasing n broadens out the pattern and suppresses the secondary maxima.

15.1.2.6. *Defocusing.* The effect of defocusing away from the Fraunhofer receiving plane is given by a term of the form

$$e^{i(u/2)\rho^2} = e^{i(2\pi/\lambda)W_{20}\rho^2} \tag{15-22}$$

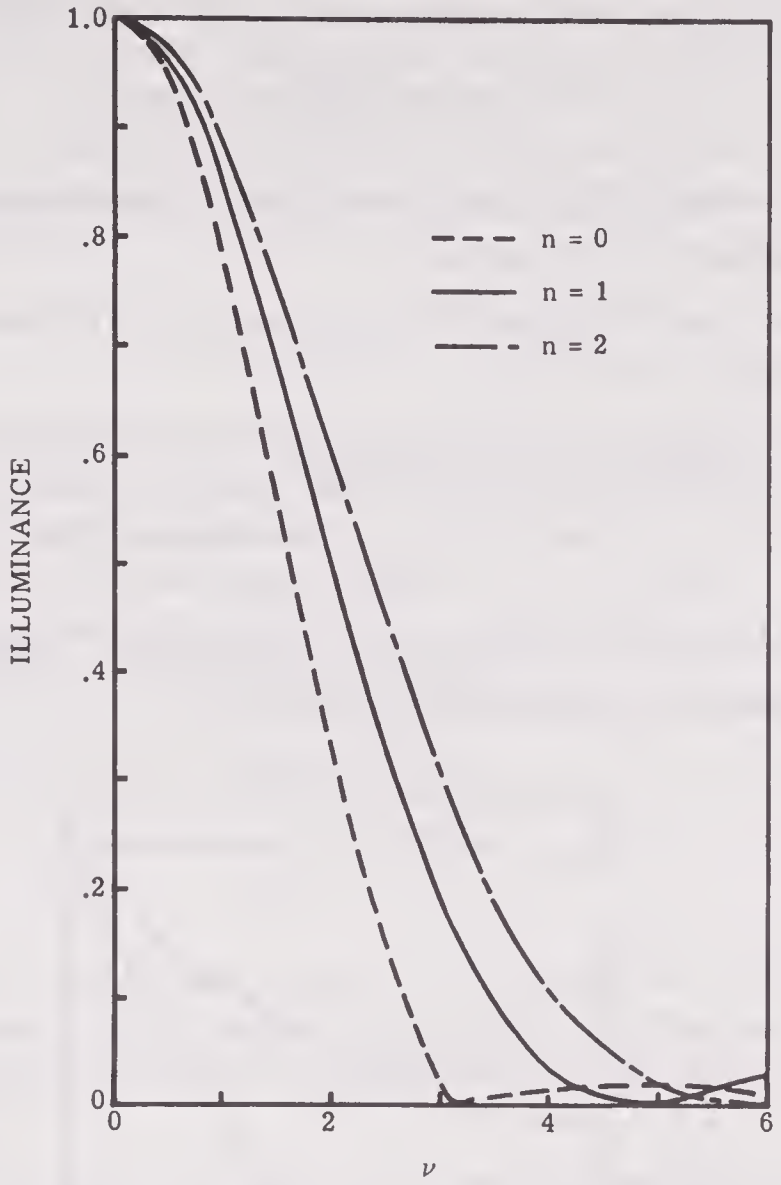


FIG. 15-5. Distribution of illuminance for Straubel pupil function; $n = 0, 1, 2$.

where W_2 is the defocusing coefficient measured in wavelength units, and u is the dimensionless defocusing parameter. In the special case of an aberration-free circular aperture the distribution of illuminance is symmetric about $W_2 = 0$. The resulting point spread function is not expressible in simple form but approaches exist which are summarized by Barakat [3]. The main effect of defocusing is to broaden the patterns.

15.1.2.7. *Rectangular and Slit Apertures.* The complex amplitude in the Fraunhofer receiving plane for a rectangular aperture is

$$a(X,Y) = \int_{-1}^{+1} A(\beta)\, e^{i\beta X}\, d\beta \int_{-1}^{+1} A(\gamma)\, e^{i\gamma Y}\, d\gamma \tag{15-23}$$

A special case of the rectangular aperture is the slit aperture, a very narrow rectangular aperture where variations in the Y direction can be neglected. Therefore,

$$a(X) = \int_{-1}^{+1} A(\beta)\, e^{i\beta X}\, d\beta \tag{15-24}$$

The center of the aperture is taken to be zero. If the limits of the aperture are $(\pm b/2)$, then

$$X = \frac{2\pi u_0 x}{\lambda} = \frac{\pi b x}{\lambda f} \tag{15-25}$$

The lossless, aberration-free Airy slit aperture, $A(\beta) = 1$, provides the illuminance

$$i(X) = \left(\frac{\sin X}{X}\right) \tag{15-26}$$

plotted in Fig. 15-1. The first zero of the slit aperture spread function is smaller than that of the corresponding circular aperture.

15.1.2.8. *Variable Pupil Functions for Slit Aperture.* Two pupil functions which can be integrated exactly are

$$A_1(\beta) = a + (1 - a)\cos\frac{\pi}{2}\beta \tag{15-27}$$

$$(0 \leqslant a \leqslant 1)$$

$$A_2(\beta) = (1 - a) + a\beta^2 \tag{15-28}$$

The behavior of the functions is shown in Fig. 15-6 and 15-7. Section 15.4 on apodization contains a discussion of additional pupil functions.

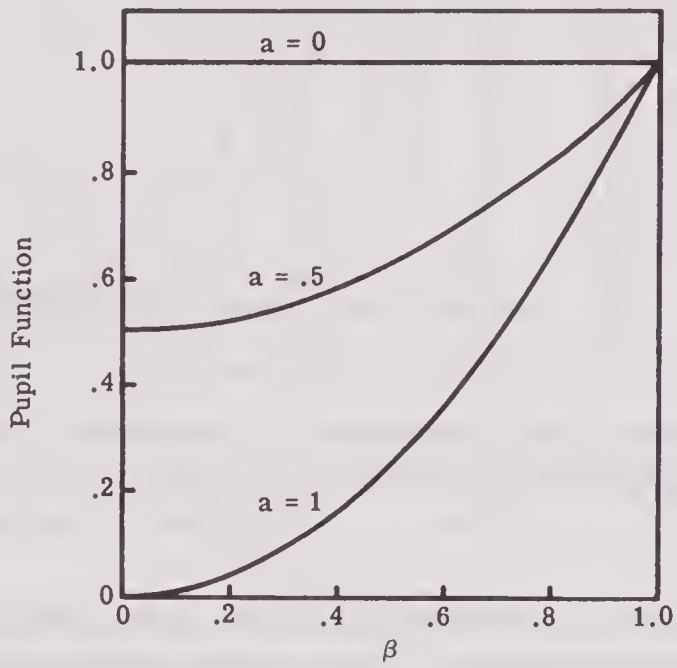


FIG. 15-6. Pupil function for slit aperture of form $A_1(\beta) = a + (1 - a)\,[\cos(\pi/2)]\beta$.

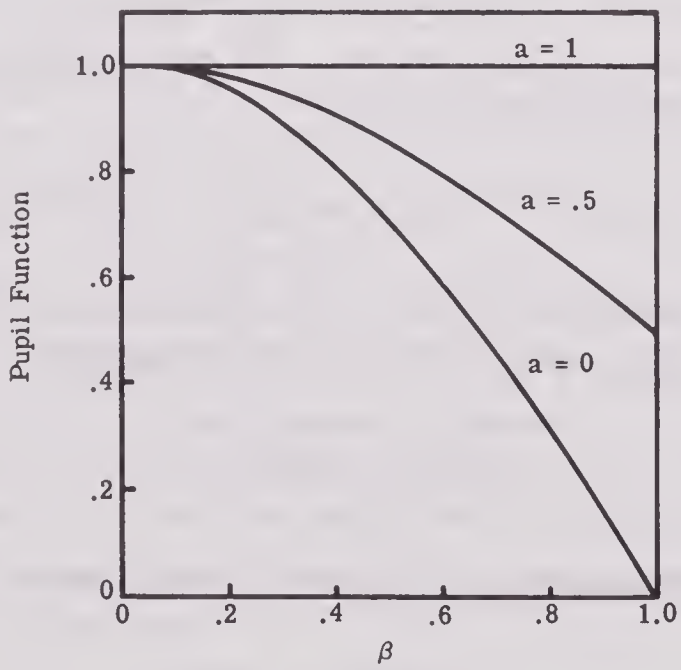


FIG. 15-7. Pupil function for slit aperture of form $A_2(\beta) = (1 - a) + a\beta^2$.

15.1.2.9. *Strehl Criterion.* The Strehl criterion (Strehl definition, Strehl intensity ratio) is not a physically measurable quantity but is a common theoretical measure of lens performance. The Strehl criterion (SC) is defined as the ratio of the central illuminance ($X = Y = 0$) of the system under consideration (with the possible inclusion of aberrations and variable pupil function) to the central illuminance of an aberration-free Airy objective; thus

$$\text{SC} = i(0, 0, A, W) = \frac{t(0, 0, A, W)}{t_{A_i}(0, 0)} \tag{15-29}$$

where A and W are included as parameters to indicate the dependence of the SC on them. The SC is unity for perfect systems. The first Luneberg apodization theorem states that the Strehl criterion is a maximum for an Airy objective ($A = 1$); any variable pupil function must decrease the Strehl criterion.

For the slit-aperture pupil functions given by Eq. (15-27) and (15-28), the SC expressions for the respective pupil functions yield:

$$i_1(0) = 1 - \frac{2}{3}a \tag{15-30}$$

$$i_2(0) = \left(1 - \frac{2}{\pi}\right) a + \frac{2}{\pi} \tag{15-31}$$

The decrease in $i(0)$ is quite appreciable for the limiting cases. If the system is aberration free, a decrease in the SC implies that the light from the central maximum must end up in the secondary maxima of the diffraction pattern because the system is normalized to constant flux.

15.1.3. Transfer Functions of Optical Systems in Incoherent Light. If $o(\sigma, \delta)$ is the distribution of intensity in the object plane, $i(X, Y)$ is the distribution of illuminance in the image plane, and $t(X, Y)$ is interpreted as the impulse response of the optical system in the spatial domain, then by the linearity hypothesis

$$i(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t(X, \sigma; Y, \delta) o(\sigma, \delta)\, d\sigma\, d\delta \tag{15-32}$$

which is the two-dimensional generalization of Eq. (15-5). Assuming spatial invariance,

$$i(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t(X-\sigma; Y-\delta) o(\sigma, \delta)\, d\sigma\, d\delta \tag{15-33}$$

A region with the property $t(X, \sigma; Y, \delta) = t(X-\sigma; Y-\delta)$ is an isoplanatic region.

Equation 15-33 is in the form of a convolution integral whose Fourier transforms obey the product theorem

$$I(\omega_x, \omega_y) = T(\omega_x, \omega_y) O(\omega_x, \omega_y) \tag{15-34}$$

where $T(\omega_x, \omega_y)$ = transfer function (frequency response) of system
$O(\omega_x, \omega_y)$ = spatial spectrum of object intensity distribution
$I(\omega_x, \omega_y)$ = spatial spectrum of image illuminance distribution
ω_x, ω_y = spatial frequencies in x and y directions (radians)

Therefore,

$$i(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} i(\omega_x,\omega_y) \exp\,[i(\omega_x X + \omega_y Y)]\, d\omega_x\, d\omega_y \tag{15-35}$$

$$I(\omega_x, \omega_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} i(X,Y) \exp\,[-i(\omega_x X + \omega_y Y)]\, dX\, dY \tag{15-36}$$

and similarly for the other Fourier transform pairs.

The fundamental equation, (15-34), is in the frequency domain and is the generalization of the classical result of monochromatic linear system theory:

$$T(\omega_x{}^0, \omega_y{}^0) = \frac{\text{Output}}{\text{Input}} \tag{15-37}$$

which is obtained by making $O(\omega_x, \omega_y)$, $I(\omega_x, \omega_y)$ behave as delta functions:

$$O(\omega_x, \omega_y) = \delta(\omega_x - \omega_x{}^0)\delta(\omega_y - \omega_y{}^0) \tag{15-38}$$

$$I(\omega_x, \omega_y) = \delta(\omega_x - \omega_x{}^0)\delta(\omega_y - \omega_y{}^0) \tag{15-39}$$

where $\omega_x{}^0$ and $\omega_y{}^0$ are constants.

In terms of the pupil function $a(X,Y)$:

$$T(\omega_x, \omega_y) = \frac{1}{T(0,0)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(\beta, \gamma)A^*(\beta-\omega_x;\, \gamma-\omega_y)\, d\beta\, d\gamma \tag{15-40}$$

$$T(0,0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |A(\beta, \gamma)|^2\, d\beta\, d\gamma$$

where $T(0,0)$ is a normalizing constant such that $|T(\omega_x, \omega_y)| \leq 1$. By this relationship the transfer function is determined by a knowledge of the pupil function, which in itself is a function of the design data of the system (*i.e.,* refractive indices, radii of curvature). The basic theorem is that the (incoherent) transfer function is the convolution of the pupil function across the exit pupil of the system. The geometry is illustrated in Fig. 15-8. This is the most practical way of determining the transfer function. Except for the simplest cases, the alternative method of computing $T(\omega_x, \omega_y)$ by evaluating the Fourier transform of $t(X,Y)$ is too complicated.

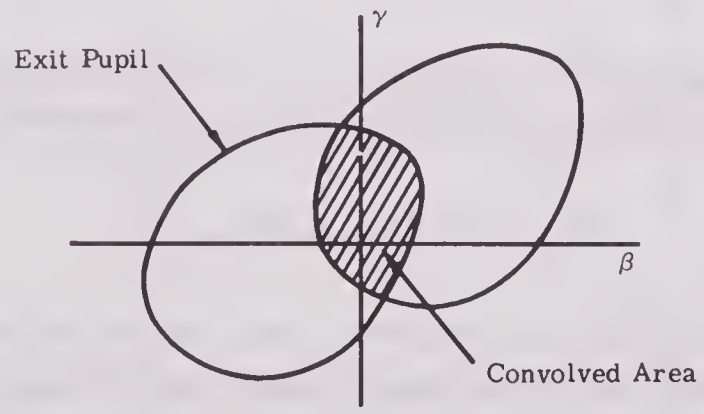


FIG. 15-8. Geometry of convolution integral required for evaluation of transfer function.

In the absence of any aberration ($W = 0$) and any apodization ($A = 1$), the transfer function reduces to the ratio of two areas. This allows one to compute the transfer function of a perfect Airy system with an arbitrary aperture by analog methods or a planimeter.

At some point (ω'_x, ω'_y) the convolved area will be zero and $T(\omega_x, \omega_y)$ remains zero for any values exceeding ω'_x, ω'_y. An optical system acts as a low-pass filter in the spatial domain, and has a sharp cutoff. The convolved area is a maximum for zero frequency ($\omega_x = \omega_y = 0$) and is normalized to unity for convenience.

The slope of the transfer function at the origin is independent of the presence of aberrations. This has the important effect of lessening the influence of aberrations at low spatial frequencies.

In general, $T(\omega_x, \omega_y)$ is a complex quantity

$$T(\omega_x, \omega_y) = |T(\omega_x, \omega_y)|\ e^{i\theta} \tag{15-41}$$

In the special case where the system is aberration free and has a real pupil function, the phase angle θ is zero and there is no phase shift; *i.e.*, $T(\omega_x, \omega_y)$ is real. Furthermore, the transfer function is real for any symmetric aberration (defocusing, spherical aberration). Only asymmetric aberrations lead to complex-valued transfer functions.

15.1.3.1. *Transfer Function as Contrast Ratio.* Define image contrast by:

$$C_I = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \tag{15-42}$$

A sine wave test target has an intensity distribution of the form

$$O(\omega_x{}^0, 0) = A_0 + B_0 \cos(\omega_x{}^0 X) \tag{15-43}$$

The spatial periodicity is taken in one direction only; A_0 and B_0 are constants. The object contrast is

$$C_0 = \frac{O_{max} - O_{min}}{O_{max} + O_{min}} = \frac{B_0}{A_0} \tag{15-44}$$

Thus

$$O(\omega_x{}^0, 0) = A_0[1 + C_0 \cos(\omega_x{}^0 X)] \tag{15-45}$$

and hence

$$I(\omega_x{}^0, 0) = D_0[1 + T(\omega_x{}^0, 0) C_0 \cos(\omega_x{}^0 X)] \tag{15-46}$$

Thus, the modulus of the transfer function is a measure of the ratio of the image contrast to the object contrast. The maximum value of the transfer function is at zero frequency (dc response), and it is standard procedure to adopt the normalization that $T(\omega_x, \omega_y) = 1$ at the origin, $\omega_x = \omega_y = 0$. The phase θ of the transfer function is a measure of the amount of lateral displacement of the image from the geometric center. It is possible to measure θ experimentally but not as easily as measuring the modulus. A linear phase shift translates the diffraction image and thus has no detrimental effect. Nonlinear phase shifts introduce harmonic distortion which, of course, results in serious deterioration of the image. For example, coma introduces a cubic phase shift.

15.1.3.2. *Strehl Criterion* (SC) *Expressed in Terms of Transfer Function.*

$$\mathrm{SC} = \frac{t(0,0)}{T_{A_i}(0,0)} = \left| \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} t(\omega_x, \omega_y)\, d\omega_x\, d\omega_y}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} T_{A_i}(\omega_x, \omega_y)\, d\omega_x\, d\omega_y} \right| \tag{15-47}$$

upon setting $X = Y = 0$ in Eq. (15-33). The SC is the area under the transfer function. This is an alternative method of obtaining the SC and is convenient if only the transfer function is known. The maximum SC (unity) is given by an aberration-free Airy objective ($A = 1$). Thus, the area under the transfer function of an actual system can be greater than that of a perfect system over a limited range; however, the overall response cannot be greater.

A commonly accepted tolerance condition is that the SC shall not fall below 0.8. A system that satisfies this criterion gives an image which is only slightly inferior to that of a perfect system.

15.1.3.3. *Transfer Function Expressions.* For a slit aperture of length b and a constant pupil function

$$T(\Omega) = \frac{1}{T(0)} \int_{-1}^{1} A(\beta)A^*(\beta - \omega_x)\, d\beta = \frac{1}{T(0)} \int_{(\omega_x/\mu_0)-1}^{1} d\beta \tag{15-48}$$

The integration is only over the intersection and the pupil function is constant, and ω_x must be normalized to μ_o as is β, then

$$T(\Omega) = \frac{1}{T(0)}(1 - \Omega) \tag{15-49}$$

Here

$$\Omega = \frac{\omega_x}{2\mu_0} = \frac{2\pi f_x}{b/2f} = \frac{4\pi F}{\lambda_x} = 2\omega_x F \tag{15-50}$$

where F is the focal ratio and $f_x = 1/\lambda_x$.

When $\Omega = 1$, the transfer function vanishes and the cutoff frequency is

$$\omega_x' = 1/2F \tag{15-51}$$

15.1.3.4. *Transfer Function for Circular Aperture.* For a circular aperture of radius b with a constant pupil function,

$$T(\Omega) = \frac{2}{\pi}\left[\cos^{-1}\Omega - \Omega(1 - \Omega^2)^{1/2}\right] \quad (0 \leqslant \Omega \leqslant 1) \tag{15-52}$$

where Ω is given by Eq. (15-50) with the diameter of the circle replacing the length of the slit. The transfer function for the circular aperture is not a straight line like that of the slit aperture. Figure 15-9 shows the behavior of the transfer functions for slit and circular apertures. Table 15-1 gives the values for $T(\Omega)$ for different values of Ω ($0 \leqslant \Omega \leqslant 1$).

15.1.3.5. *Transfer Function for Annular and Annulus Apertures.* The curves shown in Fig. 15-10 are adapted from O'Neill [4]. As ϵ the obscuration ratio increases, the low-frequency response decreases with a corresponding increase in the high-frequency response. In fact, as ϵ approaches unity, the transfer function (being the convolution of two very thin rings) will have a spike of height 1 at the origin and a spike of height 1/2 at $\Omega = 1$. See also [5].

The transfer function for the corresponding annulus aperture is illustrated in Fig. 15-11. Whereas the annular aperture emphasizes the high frequencies, the annulus aperture emphasizes (by proper adjustment of ϵ and ϵ') the intermediate-frequency region.

15.1.3.6. *Transfer Function of a Typical Reflecting System.* When a catadioptric system is used for image formation, the distribution of illuminance of the system takes

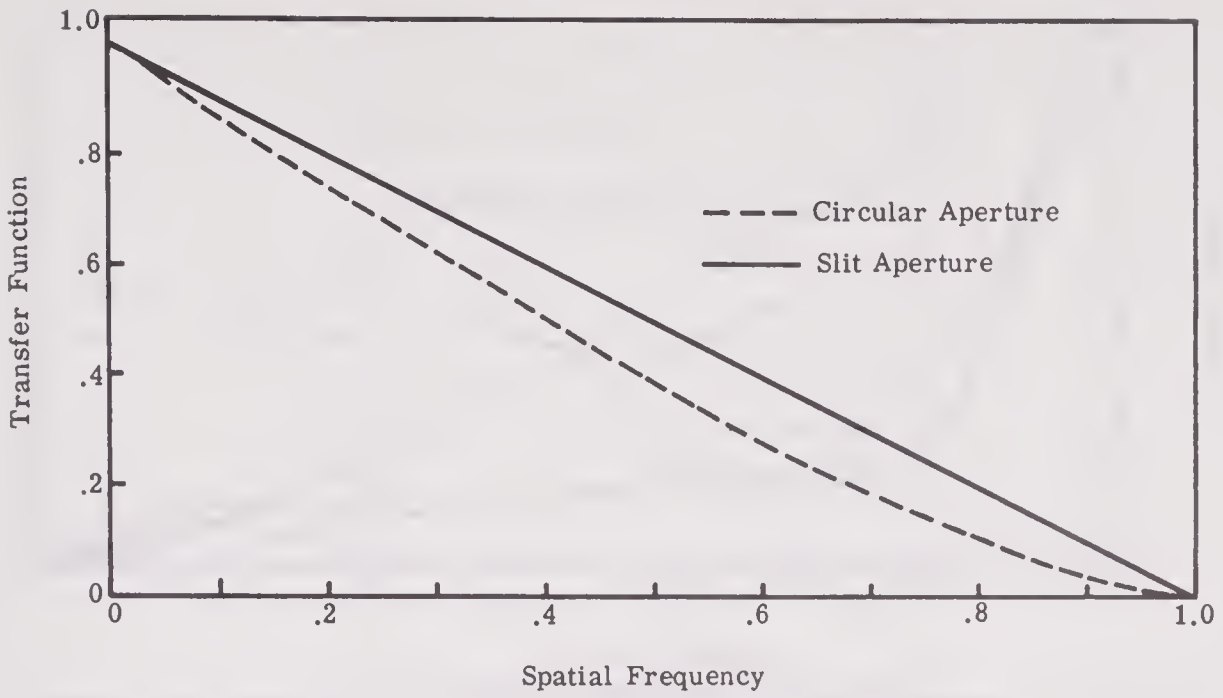


FIG. 15-9. Transfer functions for aberration-free slit and circular apertures.

TABLE 15-1. TRANSFER FUNCTION FOR ABERRATION-FREE CIRCULAR APERTURE

Ω	$T(\Omega)$	Ω	$T(\Omega)$
0	1.0000	0.55	0.3368
0.05	0.9364	0.60	0.2848
0.10	0.8729	0.65	0.2351
0.15	0.8097	0.70	0.1881
0.20	0.7471	0.75	0.1443
0.25	0.6850	0.80	0.1041
0.30	0.6238	0.85	0.0681
0.35	0.5636	0.90	0.0374
0.40	0.5046	0.95	0.0133
0.45	0.4470	1.00	0
0.50	0.3910		

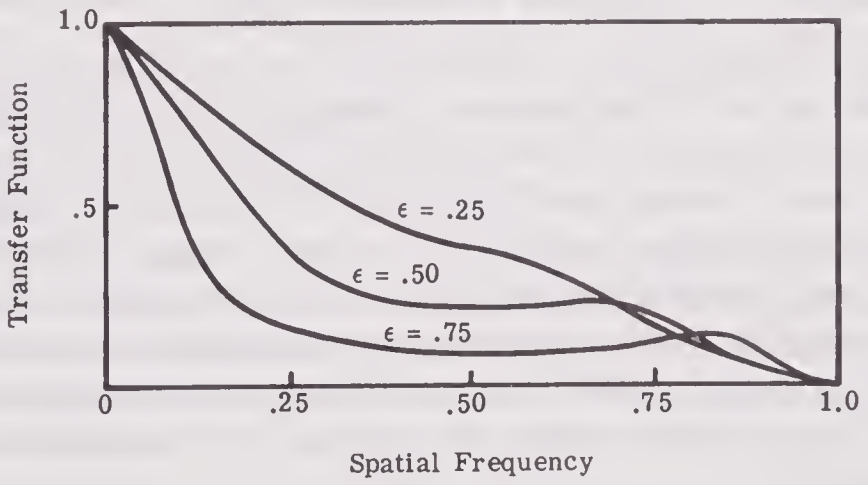


FIG. 15-10. Transfer function of annular aperture for different amounts of central obscuration [6].

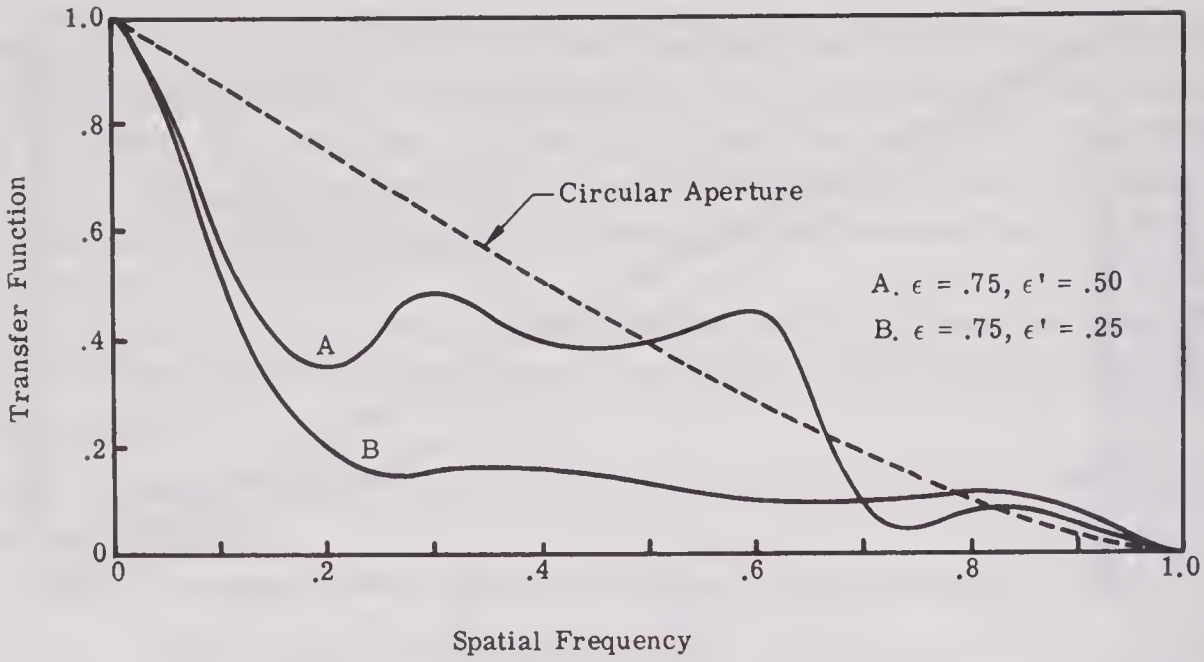


FIG. 15-11. Transfer function of annulus aperture for different ring widths.

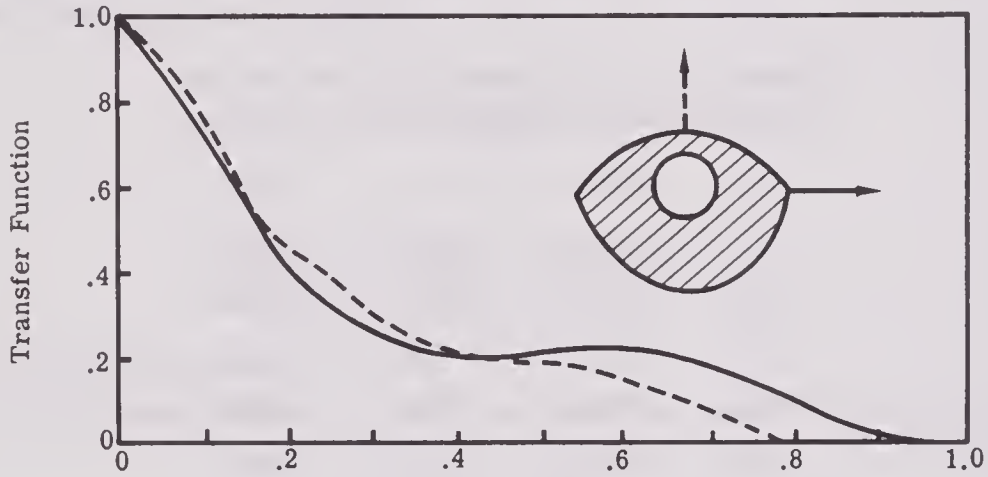


FIG. 15-12. Transfer function of a typical reflecting system as seen off-axis.

an extremely complex form due to the obscuration of the aperture, and is decidedly off-axis. The maximum response will occur when there are no aberrations; this is the true aperture limit of the system because aberrations only lower the response. Figure 15-12 shows the off-axis response of a typical catadioptric system. The transfer function is now a function of two spatial frequencies since symmetry no longer exists.

15.1.3.7. *Coherent Illumination.* Coherently illuminated situations do not usually appear in nature and can only be achieved in microscopy or specially designed laboratory equipment. The basic equation is of the form of Eq. (15-34) where the object and image spatial spectra are now of amplitude rather than of illuminance (amplitude squared). The amplitude transfer function relating the output (image) amplitude spatial spectrum to the input (object) amplitude spatial spectrum is the pupil function of the system. In Fig. 15-13, the transfer function of a slit aperture is shown for both coherent (axial illumination) and incoherent light. By displacing the point source off-axis (in the coherent case) it is possible to increase the resolution for periodic structures (Abbé's theorem); see [7] and [8].

15.1.3.8. *Cascaded Systems.* So far only the transfer function of the optical system has been employed; no provision has been made for including the effect of film or

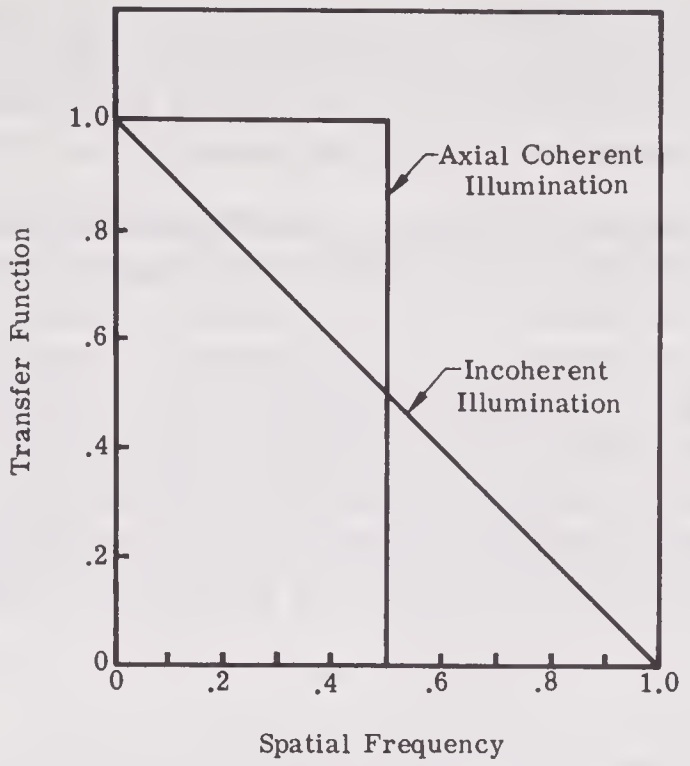


FIG. 15-13. Transfer function of aberration-free slit aperture for coherent and incoherent illumination.

detectors on the image. If the Fourier approach is used, the transfer functions of cascaded systems multiply in frequency space. Thus, the atmosphere's transfer function is T_1, that of the lens is T_2, and that of the film or detector is T_3. Now Eq. (15-34) generalizes to

$$I(\omega_x, \omega_y) = T_1(\omega_x, \omega_y)T_2(\omega_x, \omega_y)T_3(\omega_x, \omega_y)O(\omega_x, \omega_y) \tag{15-53}$$

If we neglect the atmosphere transfer function, the film or detector transfer function can materially alter the aerial image as obtained by the lens. The transfer function of films varies widely, but the use of a lens-film combination involves no new principles. The importance lies in the transfer function of the optical system. Unfortunately the transfer functions of individual lenses do not multiply to give the overall transfer function of the lens system, as they are now defined.

15.2. Resolution and Its Ramifications

Although resolution criteria exist for special situations, these criteria can only be interpreted within their limited context. Two criteria (Rayleigh and Sparrow) have been formulated specifically for dealing with point sources. The third criterion (sine-wave resolution) is for situations where sine-wave (in intensity) targets are viewed. Although other situations exist, these criteria are valuable provided that proper caution is exercised in their use and interpretation.

15.2.1. Resolution Criteria for Point Sources.

15.2.1.1. *Rayleigh Resolution Criterion.* The Rayleigh criterion states that two point sources are resolvable when the maximum of the illuminance produced by the first point source falls on the minimum of the illuminance produced by the second point source. The Rayleigh criterion is tacitly based upon two assumptions that severely restrict the generalization of the criterion: (*a*) point sources are incoherent; (*b*) point sources are of equal intensity. The two (incoherent) point sources are to be placed a distance δ from the center line of the optical system (the distance between the points is 2δ). When the critical value of 2δ is reached the value of 2δ is the (Rayleigh) limit of resolution, called δ_0. According to the Rayleigh criterion, the limit of resolution is

given by $\delta_0 = 3.832$ for the circular aperture and $\delta_0 = 3.142$ for the slit aperture. Thus a slit aperture whose length is equal to the diameter of a circular aperture has a smaller Rayleigh resolution (approximately 18%) than that of the circular aperture. In most cases the loss of light in using the slit aperture is enough to nullify the gain in resolution.

15.2.1.2. *Sparrow Resolution Criterion.* Sparrow notes that at the Rayleigh limit there is still a central minimum which lies below the adjacent maxima of the diffraction pattern. As the distance between the point sources is further decreased the central minimum will become shallower and finally disappear. Sparrow defines the limit of resolution as that value of 2δ (again defined by δ_0) for which this phenomenon takes place. As the resultant distribution of illuminance is symmetric about the origin, all the odd derivatives with respect to the lateral-displacement parameter vanish at the origin, and the analytical statement of the Sparrow resolution criterion is that the second derivative of the total distribution of illuminance vanishes on-axis; that is

$$\frac{\partial^2}{\partial v^2} i(v, \delta) = 0 \quad (v = 0) \tag{15-54}$$

which states that at zero visibility (v) the resultant distribution of illuminance undergoes no change in slope. The Sparrow limit of resolution is given by the solution of this equation. Unlike the Rayleigh criterion, the Sparrow criterion can be applied to sources whose intensities are not equal. In addition, the Sparrow criterion depends upon the coherence of the source. The resolution limits for circular and slit apertures are listed in Table 15-2. The coherent values of δ_0 are larger than the corresponding values for incoherent illumination.

TABLE 15-2. RESOLUTION LIMITS FOR AIRY-TYPE OBJECTIVES WITH POINT SOURCES OF EQUAL INTENSITY [4]

	Circular	*Slit*
Sparrow Coherent	$\delta_0 = 4.600$	$\delta_0 = 4.164$
Rayleigh	$\delta_0 = 3.832$	$\delta_0 = 3.142$
Sparrow Incoherent	$\delta_0 = 2.976$	$\delta_0 = 2.606$

It is also possible to formulate the Sparrow criterion in the spatial frequency domain. For example the slit aperture case with incoherent point sources leads to

$$\int_0^1 \Omega^2 T(\Omega, W, A) \cos(\delta_0 \Omega)\, d\Omega = 0 \tag{15-55}$$

where T is the transfer function (with the possible inclusion of aberrations W and variable pupil function A). Solutions of this equation yield the incoherent Sparrow limit δ_0. The quadratic dependence of Ω in Eq. (15-55) implies that the high-frequency components of the spatial frequency domain are the determining factors for the Sparrow incoherent criterion. Thus, any transfer function which enhances the high-frequency region (such as the annular aperture) will decrease δ_0.

15.2.2. Sine-Wave Resolution. With an incoherently illuminated sine-wave target of spatial frequency $\omega_x{}^0$ in place of the point-source objects, the distribution of illuminance in the image plane is

$$I(\omega_x{}^0, 0) = D'_0\{1 + T(\omega_x{}^0, 0)\ [\exp i\theta(\omega_x{}^0)]\ C_0 \cos(\omega_x{}^0 X)\} \tag{15-56}$$

The diffraction image of the sine-wave target will possess no gradation of illuminance when the cosine term of Eq. (15-56) vanishes, that is, when

$$\omega_x{}^0 X = \lambda/2 \tag{15-57}$$

which leads to a sine-wave cutoff frequency given by Eq. (15-51). This is the aperture-limited resolution of the system and is the maximum value that the particular system can have.

15.2.2.1. *Spurious Resolution.* When the transfer function vanishes, the illuminance in the image is uniform for spatial frequencies smaller than the cutoff frequency. The transfer functions discussed in Section 15.1 are for aberration-free systems. The inclusion of aberrations will cause the transfer function in many instances to assume the form given in Fig. 15-14, which is the transfer function for a slit aperture with one wave of third-order spherical aberration in the marginal receiving plane. This curve is typical of what is met in practice. The transfer function has regions where it takes on negative values. However, since the transfer function represents contrast, which is a positive quantity, negative values are interpreted as arising from a phase shift of amount λ. This implies that, in the regions where the transfer function is negative, black and white lines reverse their original position. This effect is called spurious resolution and is very serious because it sets an upper limit on the spatial frequencies which are useful.

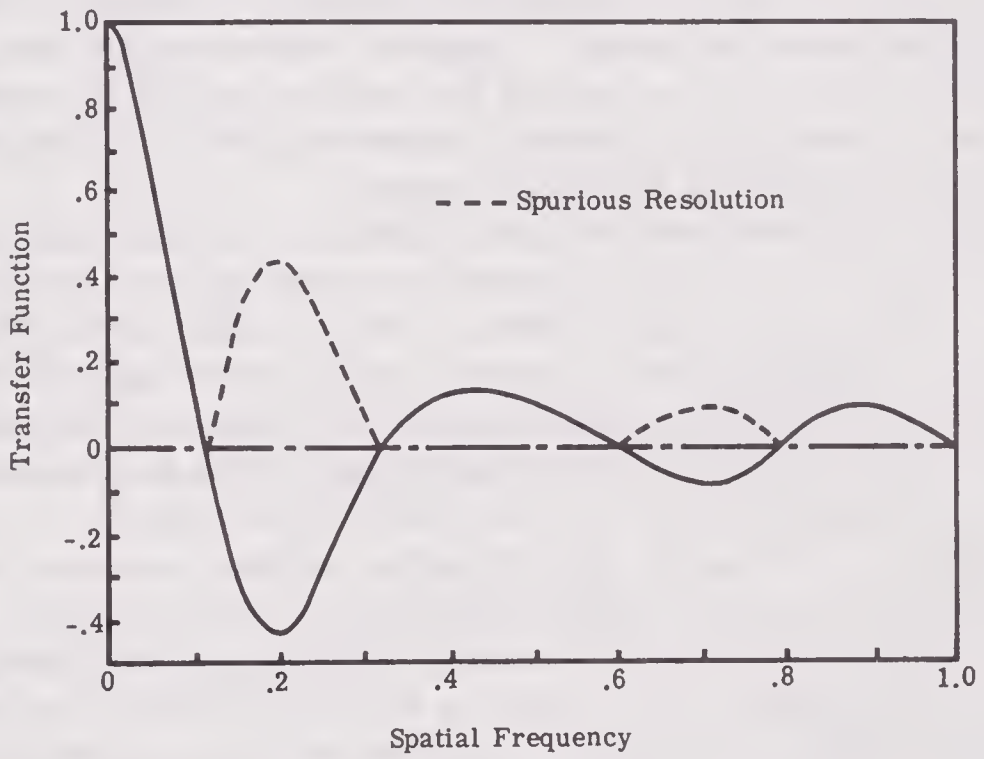


FIG. 15-14. Transfer function illustrating spurious resolution.

15.3. Effect of Aberrations on Transfer Function

15.3.1. Spherical Aberration and Coma. If an optical system is perfect, the incident spherical wave front must emerge as a spherical wave front after passing through the system. Assume that the system is rotationally symmetric, *i.e.,* has a circular aperture. The deviation from the ideal spherical wave front is measured in terms of the aberration function W. When W is identically zero, the wave-front aberrations vanish and the wave front is spherical. Wave theory of aberrations is covered extensively in [9], [10], and [11].

The aberration function for a circular aperture depends upon ρ, ϕ (polar coordinates in the exit pupil), and also upon σ (the normalized field variable). When the point

source illuminating the aperture is on the axis of the optical system, σ vanishes. Because of rotational symmetry, the variables ρ, ϕ, and σ in W can only occur in the combinations ρ^2, σ^2, and $\rho\sigma \cos \phi$. The aberration function has an expansion of the form

$$W(\rho, \phi, \sigma) = W_{20}\rho^2 + W_{20}\rho^4 + {}_1W_{31}\sigma\rho^3 \cos \phi + {}_2W_{20}\rho^2\sigma^2 + {}_2W_{22}\rho^2 \cos^2 \phi + {}_3W_{11}\sigma^3\rho \cos \phi + \text{higher-order terms} \tag{15-58}$$

where (following Hopkins notation [10] the various W coefficients represent the Seidel or third-order aberrations. The first term ($W_{20} \equiv W_2$) represents longitudinal focal shift (defocusing) and is not an actual aberration. The Seidel aberrations are

(*a*) $W_{40} \equiv W_4 =$ spherical aberration

(*b*) ${}_1W_{31} =$ coma

(*c*) ${}_2W_{20} =$ image curvature

(*d*) ${}_2W_{22} =$ astigmatism

(*e*) ${}_3W_{11} =$ distortion

All of these coefficients are in wavelength units. Spherical aberration is independent of σ and does not vanish for a point source on-axis. The third-order coma term is of great importance as it largely determines the quality of the diffraction image in the outer parts of the field. Coma can be regarded as spherical aberration for object points lying off the optical axis of the system. Spherical aberration and coma are the most serious Seidel aberrations in the sense that they are limiting. If the system has a poor transfer function on-axis (due to spherical aberration) then it will be worse off-axis. Coma is the limiting aberration for off-axis points.

The expansion of the aberration function is actually an infinite series in ρ, ϕ, σ, and the only justification for truncating the series is the tacit assumption that the aperture and field are small enough so that powers of these variables higher than the fourth can be neglected. In many practical systems such is not the case and the fifth-order aberrations must be taken into account to make the analysis of the system realistic. Analysis of the single aberrations is a necessary step in the development of transfer-function theory as applied to actual systems. See [6] and [10].

15.3.1.1. *Defocusing.* When W_2 (the defocusing coefficient measured in wavelength units) is zero, then the image is located in the paraxial receiving plane. The transfer function is an even function of W_2. The transfer-function curves for small amounts of defocusing are shown in Fig. 15-15. Note the very rapid deterioration of the contrast as the defocusing coefficient is increased. Spurious resolution appears for even as small an amount of defocusing as one wave.

The transfer function for the rectangular aperture suffering defocusing can be evaluated explicitly:

$$T(\Omega_x, \Omega_y) = R(\Omega_x)T(\Omega_y) \tag{15-59}$$

where

$$T(\Omega) = (1 - \Omega) \operatorname{sinc} [8\pi W_2\Omega(1 - \Omega)] \quad (0 \leq \Omega \leq 1) \tag{15-60}$$

The defocused transfer function for a slit aperture (very narrow rectangular aperture) is shown in Fig. 15-16 for the same values of W_2 as those corresponding to the circular aperture. The chief difference (for corresponding values of W_2) is that the changes in the slit-aperture transfer function are more pronounced than those for the circular aperture. This is a general rule for all aberrations and is due to the smoothing action

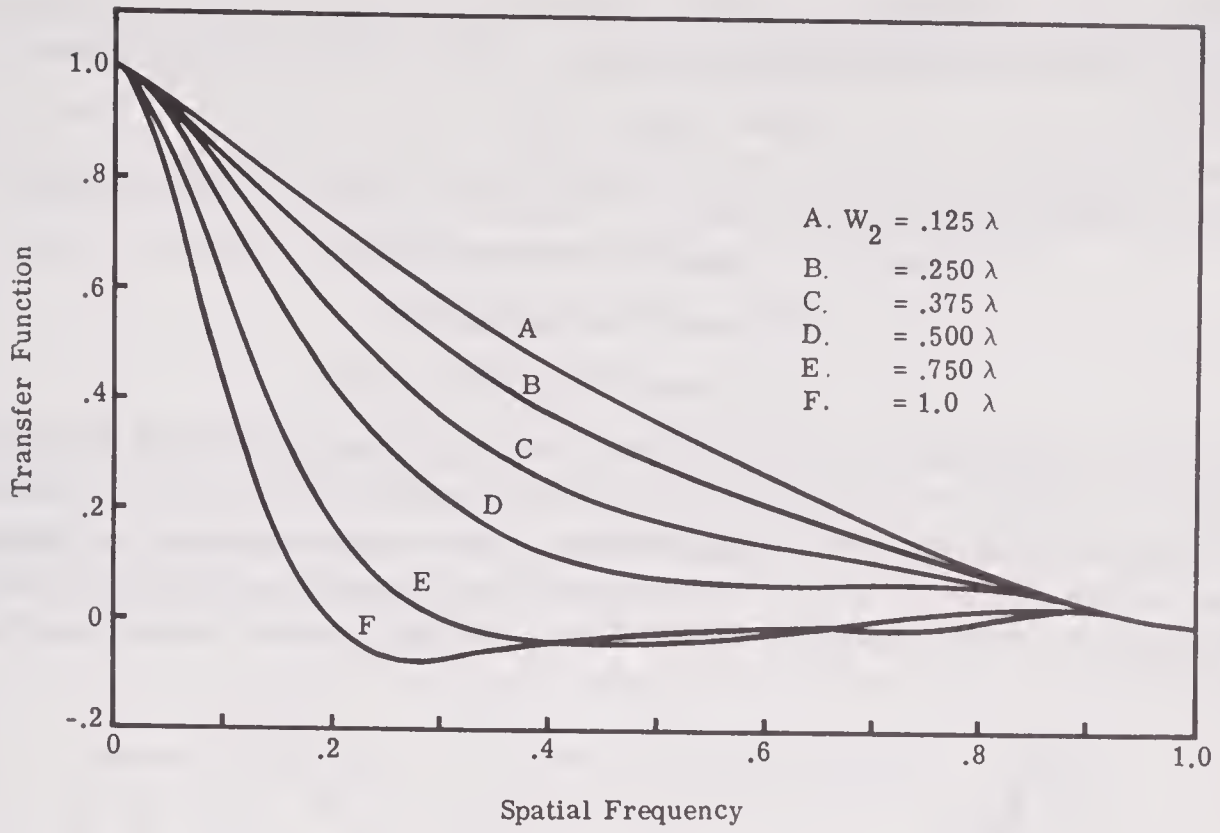


FIG. 15-15. Transfer function of circular aperture for various amounts of defocusing.

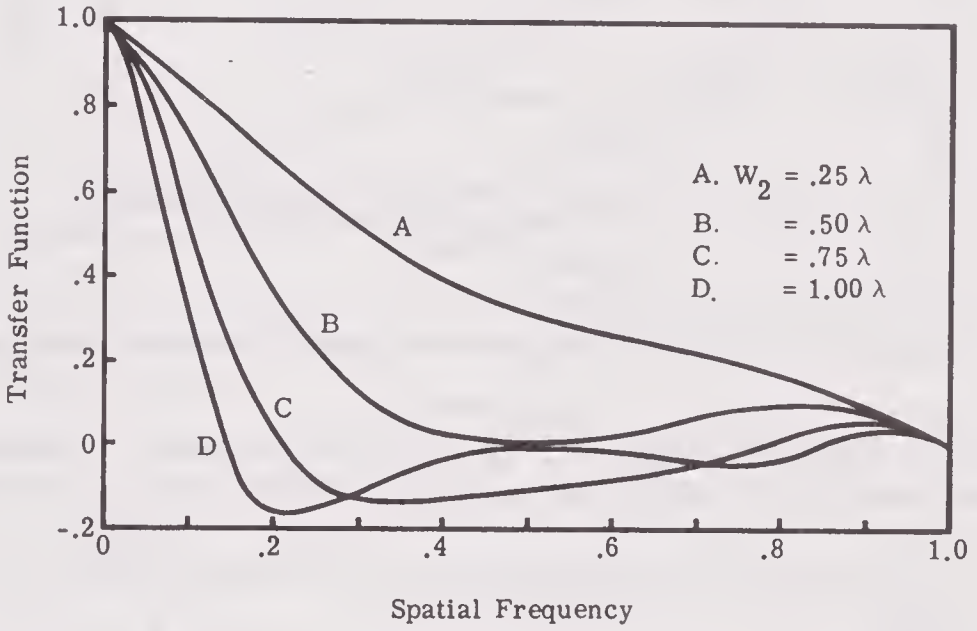


FIG. 15-16. Transfer function of slit aperture for various amounts of defocusing.

of the convolution integral which is more pronounced in two dimensions than in one dimension. See [12] and [13].

15.3.1.2. *Spherical Aberration.* A major effort has gone into determining the effects of various orders of spherical aberration on the transfer function. Unfortunately, the circular aperture is not amenable to analytic methods because of its complexity, and numerical techniques are necessary; see [14]. The Gauss quadrature method was employed by Barakat [12] for the case where the aberration function included terms up to thirteenth-order spherical aberration:

$$W = \sum_{n=1}^{14} W_{2n,0}\, \rho^{2n} \tag{15-61}$$

For the present, consider the terms W_{20}, W_{40}, and W_{60} only.

In dealing with third-order spherical aberration, the inclusion of defocusing can materially increase the response of the system. Thus, taking the aberration function in the form

$$W = W_4(\rho^4 - 2\mu\rho^2) \tag{15-62}$$

where $\mu = -W_2/2W_4$, by varying μ we can shift to any desired receiving plane:

$\mu = 0$ = paraxial receiving plane

$\mu = 1/2$ = central receiving plane

$\mu = 1$ = marginal receiving plane

In Fig. 15-17 and 15-18 the transfer functions for a circular aperture are shown for a half wave and a full wave of spherical aberration, respectively. The response in the central plane is vastly superior to the response in the other two planes. Although the marginal curves are worse than the corresponding paraxial curves in the medium-frequency region, in the extreme regions of low and high spatial frequencies the converse is true.

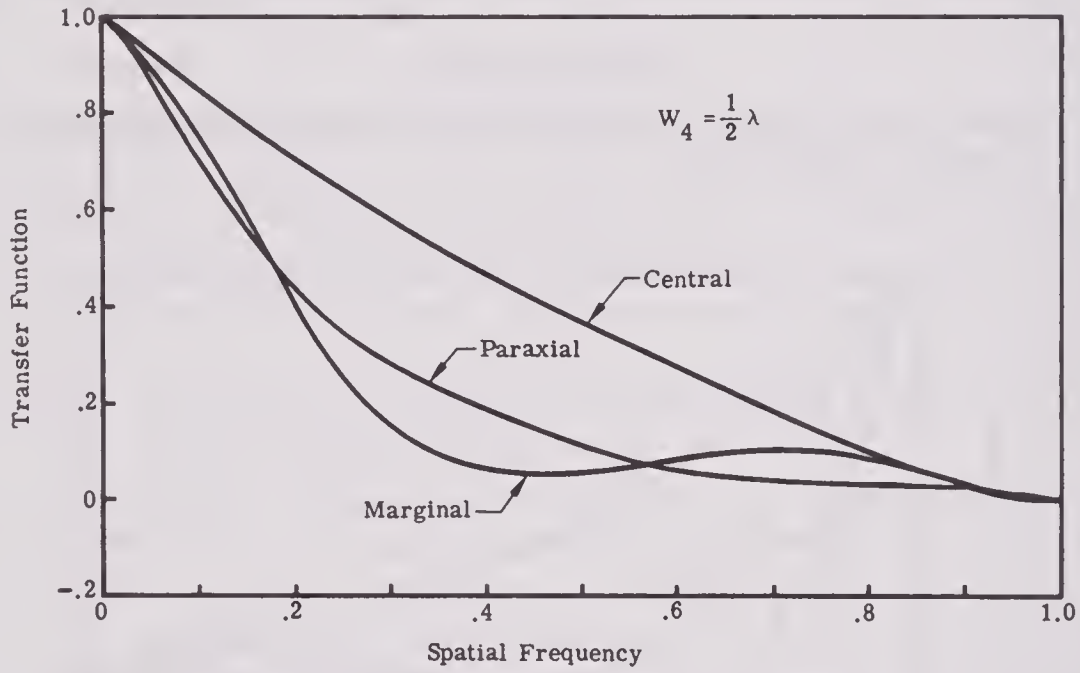


FIG. 15-17. Transfer function of circular aperture possessing a half-wave of spherical aberration in central, paraxial, and marginal receiving planes [15].

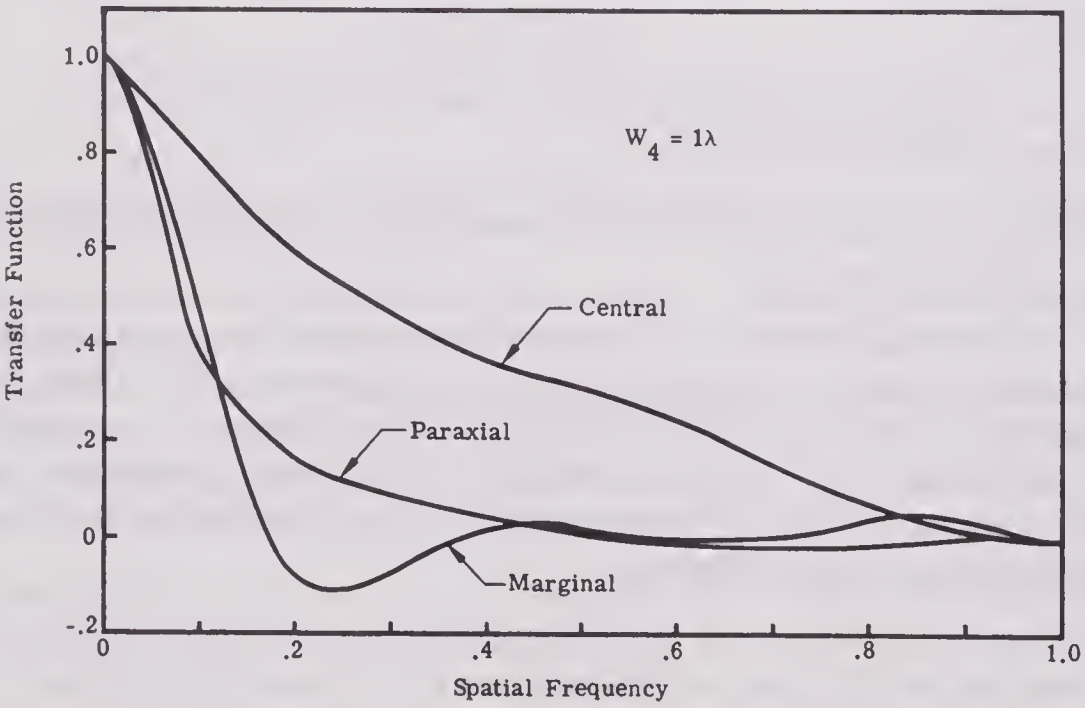


FIG. 15-18. Transfer function of circular aperture possessing one wave of spherical aberration in central, paraxial, and marginal receiving planes [15].

Parrent and Drane [16] have solved the slit-aperture problem for W_4 and W_2 using Simpson's rule. Barakat [15] has utilized Gauss quadrature, and the curves shown in Fig. 15-19 and 15-20 were evaluated on a high-speed computer. The results are similar to those for the circular aperture, although the circular aperture curves are smoother.

The extension of the analysis to include W_6 (fifth-order spherical-aberration coefficient) is straightforward, and the aberration function is of the form

$$\begin{aligned} W &= W_6\rho^6 + W_4\rho^4 + W_2\rho^2 \\ &= W_6[\rho^6 + \alpha\rho^4 + \beta\rho^2] \end{aligned} \tag{15-63}$$

The addition of the two parameters α and β presents a bewildering array of possible combinations. A theory has been developed by Marechal which is based upon setting tolerances on the wave front in terms of the Strehl criterion. This theory is useful only for fairly small aberrations.

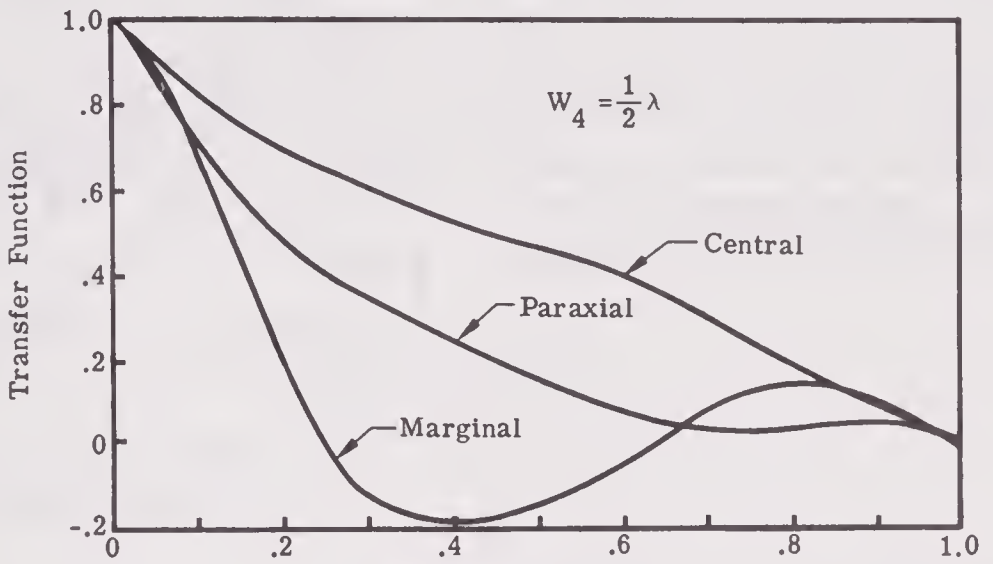


FIG. 15-19. Transfer function of slit aperture possessing a half-wave of spherical aberration in the central, paraxial, and marginal receiving planes.

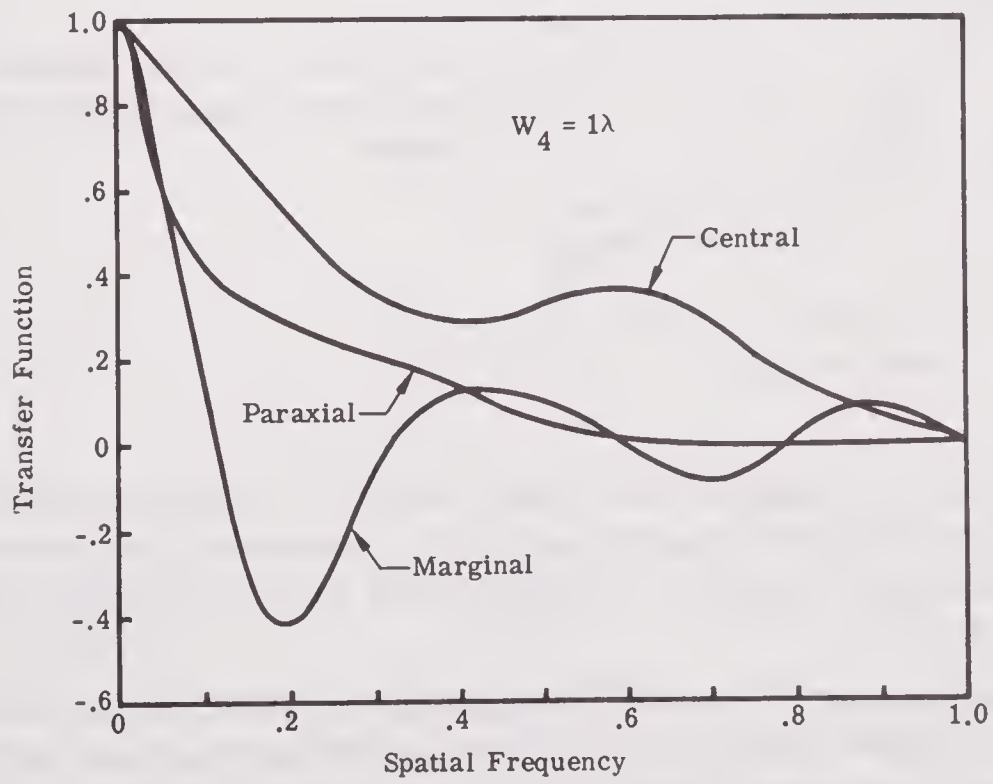


FIG. 15-20. Transfer function of slit aperture possessing one wave of spherical aberration in the central, paraxial, and marginal receiving planes.

15.3.1.3. *Coma.* Coma being an asymmetric aberration therefore has a transfer function with nonzero phase shift. Goodbody [17] has performed calculations and we quote from his work. The calculations are best performed when the aberration function is expressed in rectangular coordinates.

$$W = {}_1W_{31}(x^2 + y^2)\,(x \sin \psi + y \cos \psi) \tag{15-64}$$

Two salient points emerge from Goodbody's study:

(*a*) The paraxial focus is the best focus in the sense that defocusing results in the deterioration of the transfer function in the low-frequency region.

(*b*) The response for $\psi = 0$ is better than for any other value of ψ in the low-frequency region.

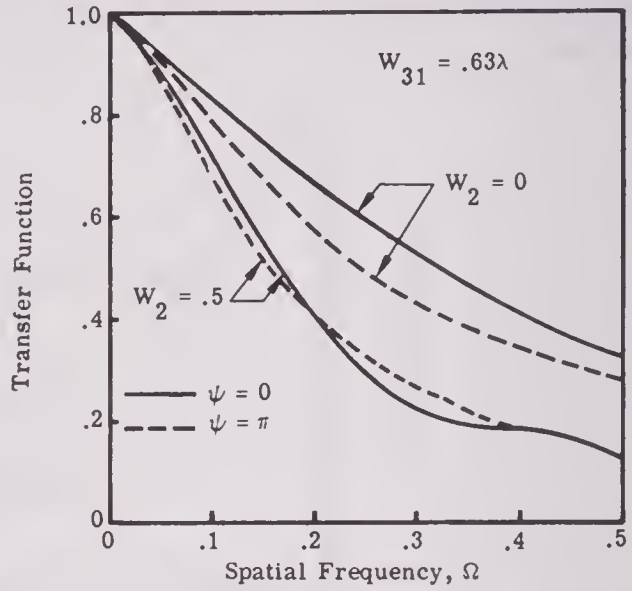


FIG. 15-21. Transfer function for third-order coma of amount 0.63 λ in two receiving planes.

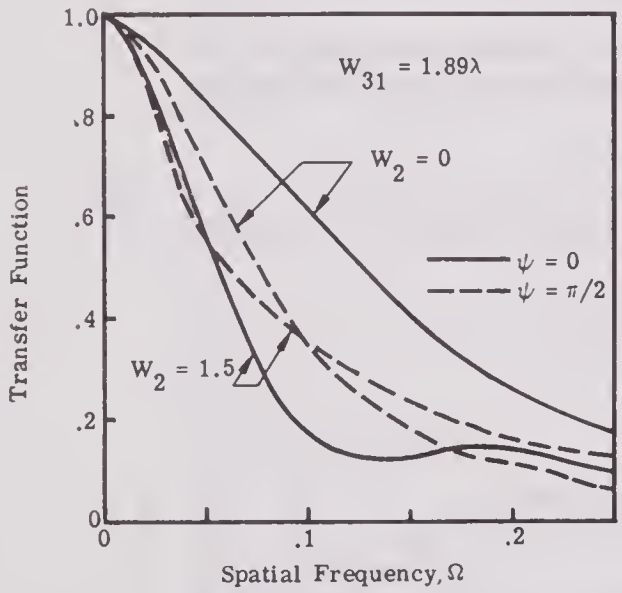


FIG. 15-22. Transfer function for third-order coma of amount 1.89 λ in two receiving planes.

The two points are illustrated in Fig. 15-21 and 15-22. The phase shift introduced by coma is nonlinear and is the cause of the harmonic distortion of the image. If the phase shift were linear with frequency, the effect would merely be a shift of the diffraction image; see [18].

15.3.2. Computation of Transfer Function of an Actual System. The fundamental problem of the transfer function from the point of lens designers is to find the functional relation between the transfer function and the design data. This problem has not been solved. However, in a given lens system, it may be necessary to obtain the transfer function directly from the ray-trace data.

This problem has been attacked by Barakat [12] and Barakat and Morello [19] using Gauss quadrature theory. Barakat's treatment is very complicated; see the actual papers for computational details and numerical results.

The schematic presented in Fig. 15-23 illustrates the sequence of basic steps. Once the wave front has been obtained it is then necessary to fit a curve by some appropriate approximation scheme, to yield a polynomial or rotational expression for W. Once the wave front is known three basic quantitites can be computed:

(1) Transfer function

(2) Distribution of illuminance

(3) Total illuminance

The distribution of illuminance due to a point source can be computed by evaluating the Kirchhoff diffraction integral Eq. (15-15). It is also of interest to know the total illumination in the various rings of the diffraction pattern, that is the fraction L of the total energy that falls within a circle of radius v_0 about the axial point in a given receiving plane. Obviously L vanishes when v_0 is zero and approaches unity as v_0 becomes infinite. Although the illuminance isophotes (lines of constant illuminance) are extremely complicated, the corresponding isophotes of total illumination are smooth

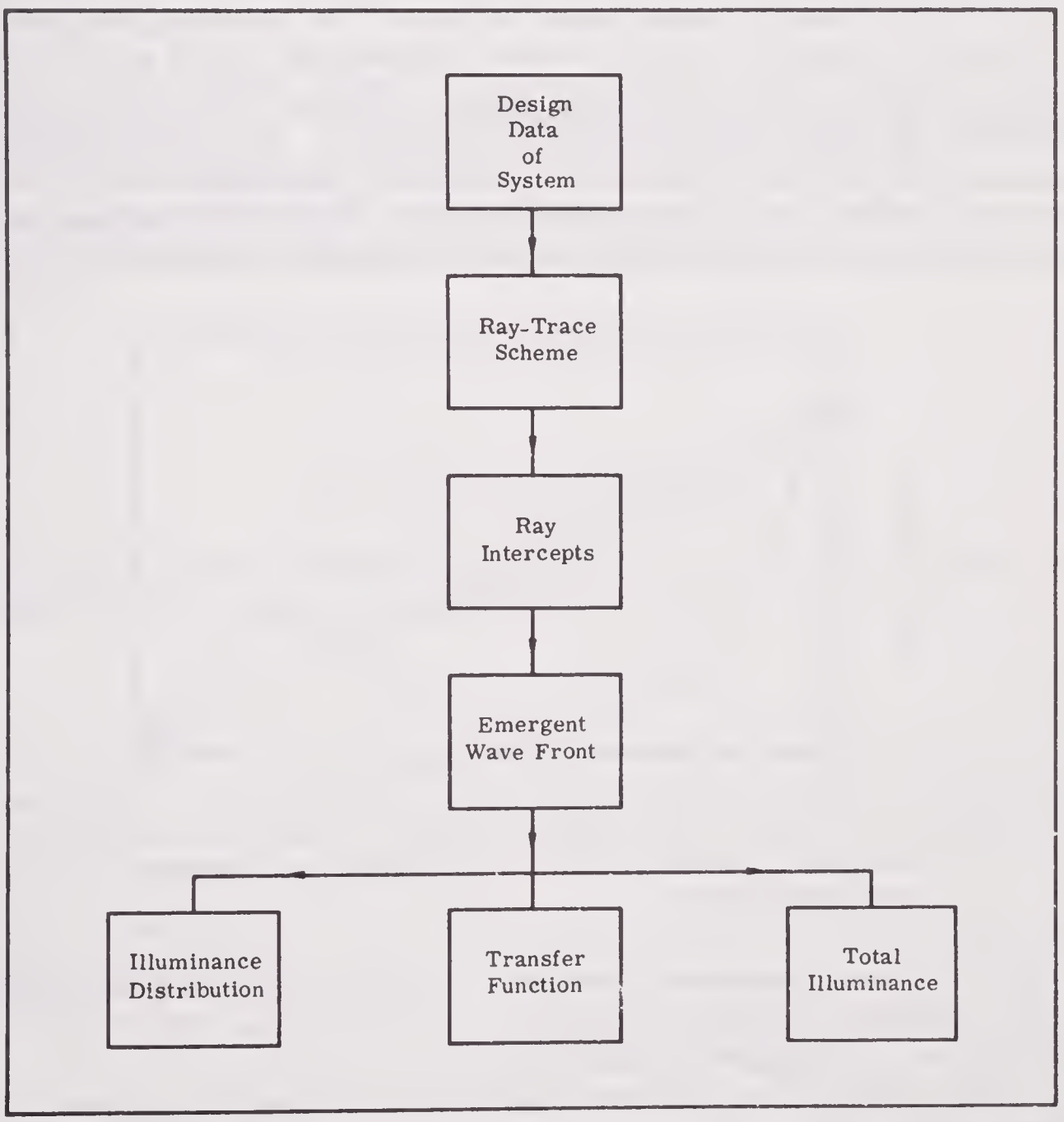


FIG. 15-23. Schematic of steps required for computation of transfer function and associated functions from design data.

functions. This is a consequence of the fact that the integration over the illuminance is a smoothing operation in the sense that the average properties of the diffraction pattern are brought to the foreground.

Other information in this area can be found in [20] and [21].

15.4. Effect of Apodization on the Transfer Function

The theory of apodization is concerned with the possibility of determining the amplitude distribution over the exit pupil (pupil function) in order to achieve some prespecified distribution of illuminance over a designated receiving plane in the image field. The usual attempts involve an expansion of the amplitude distribution over the exit pupil (pupil function) in a convenient set of functions, *e.g.*, Hermite, lambda, and Legendre.

There are two theoretical approaches to apodization. The first and simplest is to choose some pupil function and to determine the resultant transfer function, spread function, etc. Second and more difficult is the synthesis problem; that is, given a prespecified quantity, determine the required pupil function.

15.4.1. Altering High- or Low-Frequency Response by Apodization. Two important general apodization theorems are:

(1) If the pupil function weighs against the center of the aperture, then in the high-frequency region the transfer function increases over that of an Airy system.

(2) If the pupil function weighs against the edge of the aperture, then in the low-frequency region the transfer function increases over that of an Airy system.

As two examples which illustrate these theorems, consider a slit aperture with pupil the aperture; the latter against the edge of the aperture. The transfer functions for functions given by Eq. (15-27) and (15-28). The former weighs against the center of the aperture; the latter against the edge of the aperture. The transfer functions for these two pupil functions are shown in Fig. 15-24 and 15-25 for three values of a ($= 0, 0.5, 1.0$).

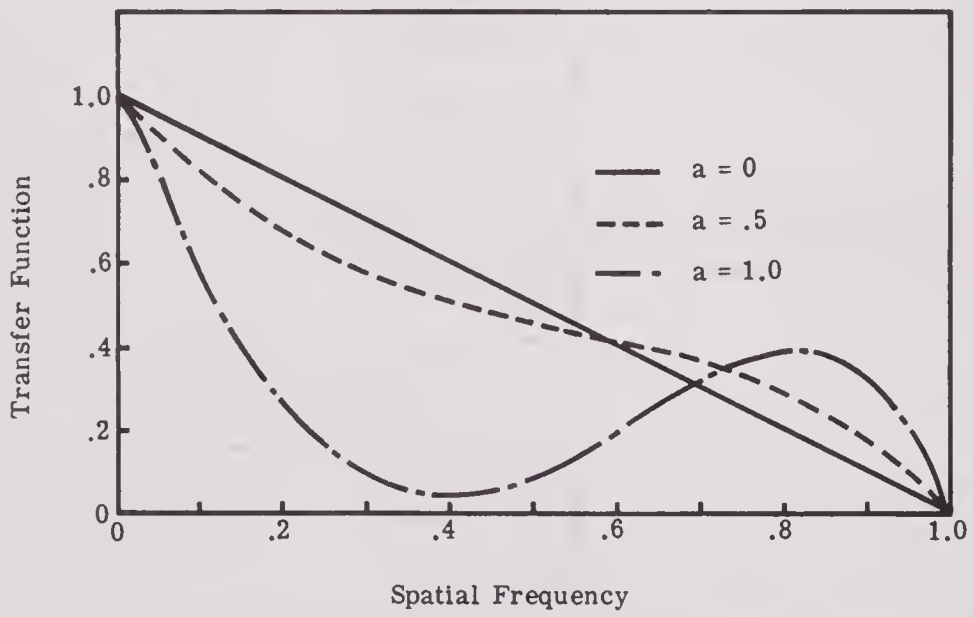


FIG. 15-24. Transfer function of apodized slit aperture having pupil function $A_1(\beta)$.

15.4.2. Luneberg Apodization Problems. Luneberg [22] formulated a series of important apodization problems, but gave the explicit solution to the first only. For a solution of the remaining problems see [23].

The most important of the theorems arises from the first Luneberg problem, which is to determine the pupil function that maximizes the central illuminance (essentially the Strehl criterion) subject to the condition that the total energy passing through the

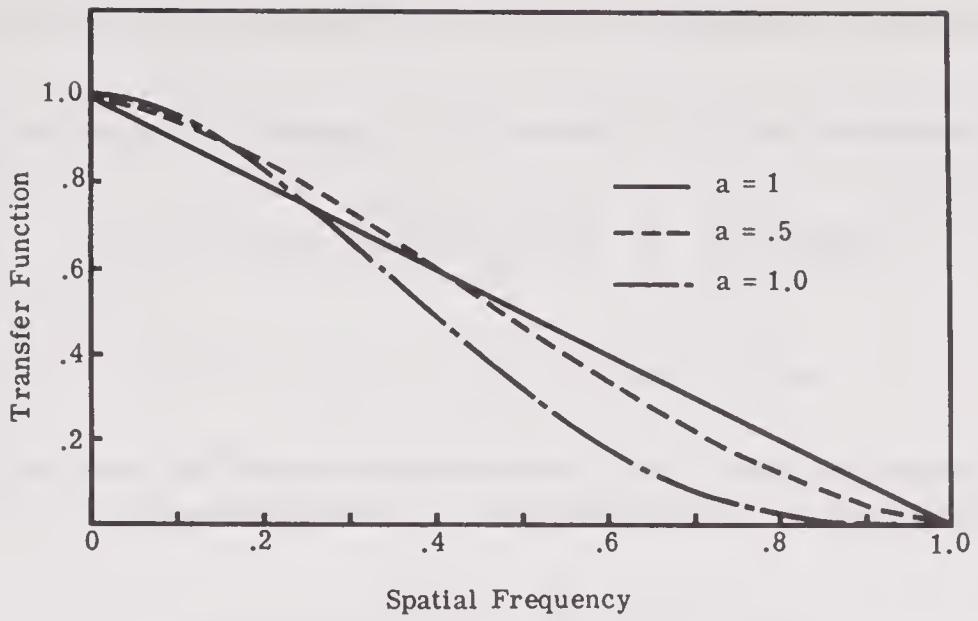


FIG. 15-25. Transfer function of apodized slit aperture having pupil function $A_2(\beta)$.

aperture is constant. The final result is the pupil function which yields the maximum Strehl criterion, *i.e.*, the Airy pupil function ($A = 1$). In other words, any apodization scheme will lower the Strehl criterion.

15.5. Merit Factors

15.5.1. Linfoot Quality Factors. The most important quality factors are the three proposed by Linfoot [24]. These factors determine on a statistical basis the degree to which an optical system reproduces various selected aspects of the object.

15.5.1.1. *Relative Structural Content.* The first quality factor (and in some respects, the most important) is the relative structural content F defined in the spatial frequency domain by

$$F = \frac{1}{C_0} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |T(\omega_x, \omega_y)|^2 O(\omega_x, \omega_y)\, d\omega_x\, d\omega_y \tag{15-65}$$

where

$$C_0 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} O(\omega_x, \omega_y)\, d\omega_x\, d\omega_y \tag{15-66}$$

and $O(\omega_x, \omega_y)$ is the power spectrum of the object. In the special but important case of a flat object spectrum [$O(\omega_x, \omega_y) = \text{constant}$], F becomes

$$F' = \frac{1}{F_0} \int\int |T(\omega_x, \omega_y)|^2\, d\omega_x\, d\omega_y \tag{15-67}$$

where F_0 is now the area of the spatial-frequency domain over which the transfer function exists. The contribution to F by the transfer function is always positive because T appears as a squared quantity. Thus the effects of spurious resolution are not accounted for using this quality factor.

15.5.1.2. *Fidelity Defect.* The second quality factor is the fidelity defect defined in the frequency domain by

$$\theta = \frac{1}{C_0} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [1 - T(\omega_x, \omega_y)]^2 O(\omega_x, \omega_y)\, d\omega_x\, d\omega_y \tag{15-68}$$

The effect of spurious resolution is now taken into account because the transfer function now enters linearly (as well as quadratically) into θ. The fidelity-defect factor is a

measure of the degree to which the maxima and minima of the object and image are superposed.

15.5.1.3. *Correlation Quality.* The correlation quality Q is given by

$$Q = \frac{1}{C_0}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} T(\omega_x, \omega_y)O(\omega_x, \omega_y)\, d\omega_x\, d\omega_y \tag{15-69}$$

$$= \frac{1}{2}(F + \theta)$$

As the second equation shows, Q is a linear combination of the first two Linfoot quality factors. In the case of an object with a flat spectrum, Q becomes

$$Q' = \frac{1}{C_0}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} T(\omega_x, \omega_y)\, d\omega_x\, d\omega_y \tag{15-70}$$

which is the Strehl criterion introduced previously. Thus, Q is a generalization of the Strehl criterion.

15.5.2. Plane of Best Focus. There is no general method for determining the plane of best focus because it depends upon the object being viewed. The investigations for three important cases are summarized:

(1) For a point source

(2) For a sine-wave target

(3) For a random-detail target

15.5.2.1. *Point Source.* The plane of best focus for a point source is that in which the Strehl criterion is a maximum. This definition gives a clear-cut answer only for small aberrations in systems with spherical aberration and defocusing. When the aberrations are large the answers are not unique and in fact the Strehl criterion is almost useless as a merit factor for determining the plane of best focus. The theory predicts that, for a system suffering third-order spherical aberration and defocusing, the plane of best focus is the central plane halfway between the paraxial and marginal planes. For values of $W_4 > 2.5\lambda$, the Strehl-Richter theory fails. The Strehl-Richter theory is closely related to the Marechal theory previously mentioned. The Marechal theory does not deal with single aberrations, but tolerances are set on the mean square value of the wave front; it requires that the mean square deformation of the wave front with reference to a spherical wave front be minimized.

15.5.2.2. *Sine-Wave Target.* The plane of best focus for periodic detail (sine-wave target) is that receiving plane for which the transfer function is a maximum at a specified frequency; *i.e.*, the plane of best focus depends upon the spatial frequency of the target. Figure 15-26 illustrates how the plane of best focus varies as a function of defocusing for a half wave of spherical aberration. Note the shift of the plane of best focus to the marginal focus ($\mu = 1$) as the spatial frequencies are increased. A similar curve for one wave of third-order spherical aberration for both slit and circular apertures is shown in [16].

15.5.2.3. *Random Detail.* In practice one does not usually encounter point-source or sine-wave targets only but rather random distributions. This case differs from the two previous ones because statistical considerations enter. As the square of the transfer function is related to the mean square fluctuations in the image, a reasonable criterion as a measure of the plane of best focus is the relative structural content F

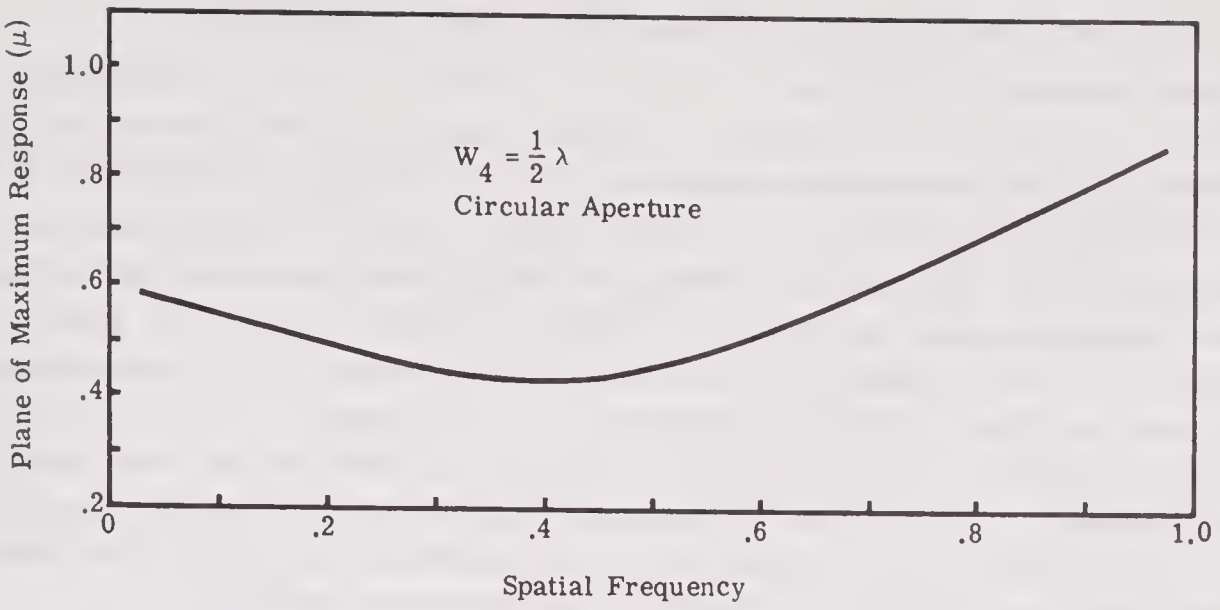


FIG. 15-26. Plane of best focus for sine wave targets.

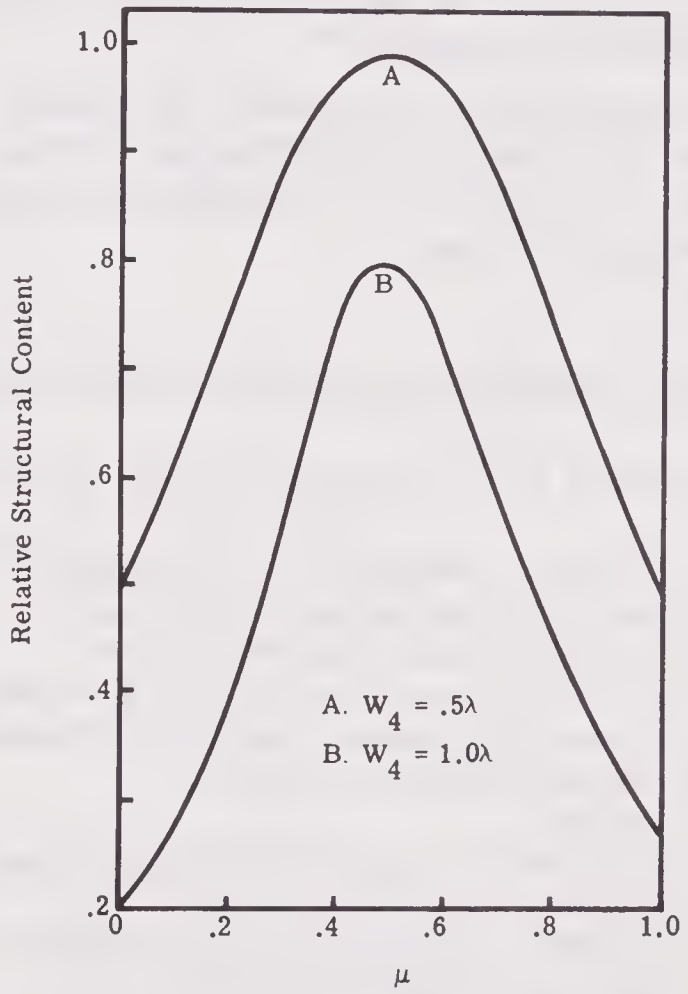


FIG. 15-27. Relative structural content of circular aperture having spherical aberration and defocusing.

defined by Eq. (15-65). The plane of best focus in the presence of random detail is that plane in which F is a maximum. The plane of best focus will depend not only upon the system through the transfer function but also upon the object via its power spectrum. To illustrate the criterion, take the case of a flat object spectrum and a circular aperture suffering third-order spherical aberration. The results of the computations are shown in Fig. 15-27 for $W_4 = 0.5\lambda$ and 1.0λ. The plane of best focus lies approximately in the central plane $\mu = 0.5$.

15.6. Frequency Response Calculations from Lens Design Data*

Frequency-response techniques have not yet reached a state of development such that they can be used for the synthesis of optical systems, although as a description of performance, and for some analysis steps, they are very useful. The Institute of Optics at the University of Rochester has a detailed design procedure programmed on a 1620 IBM computer [25]. The method is based on the standard aberration polynomial, and includes a frequency-response calculation in the procedure. Some of Dr. Hopkins' comments about frequency-response calculations are paraphrased: It is an exceedingly difficult program to write. If design requirements are stringent, wave-front errors are reduced to a minimum. If not, there is a compromise at best and not much point in the time and expense of the frequency-response calculations. Frequency response is a function of wavelength and direction; what wavelengths and directions, how many, and what weighting they should have should be applied to any average.

Frequency response can be calculated on the basis of physical optics. Miyamoto [6] discusses the calculations and their relationships in detail. Results in brief are: The region from 0 to 2λ waves of optical path difference is difficult to determine by approximate methods. As long as the geometrical image is inside the diffraction image, it is valid to use the geometrical response.

15.6.1. Computer Calculations of Frequency Response. A recent article in *Applied Optics* describes the theory for performing some of these calculations. As an example, and for possible use or adaptation, the FORTRAN frequency-response program as used by the Institute is given below.

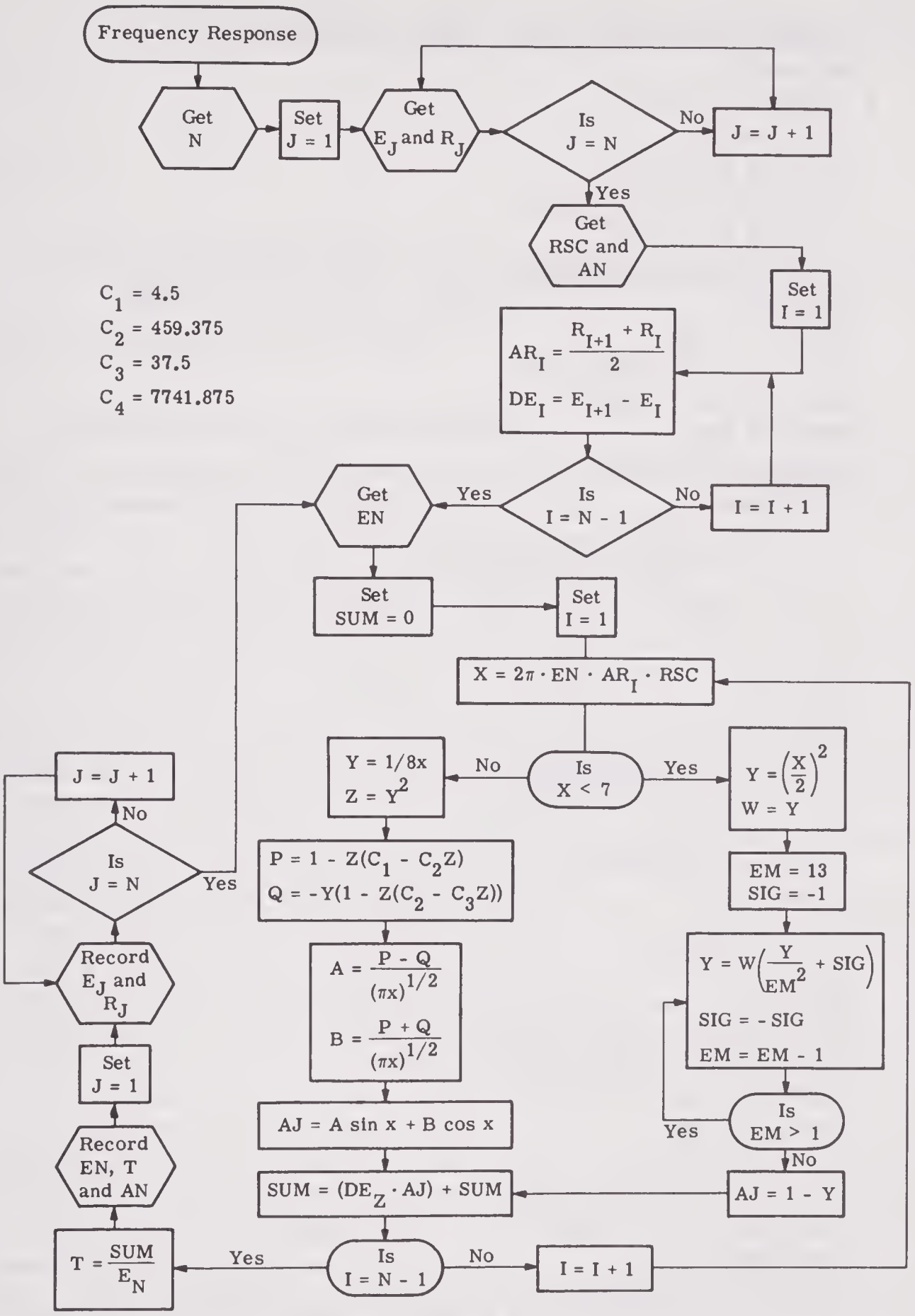
FREQUENCY-RESPONSE PROGRAM FOR THE IBM 1620

Abbreviations Used in Frequency-Response Program

R = radius of a circular area at the image
E = energy within the circle of radius R
AR = r = mean of two adjacent R values
DE = ΔE = difference of two adjacent E values
N = number of R values = number of E values
J = running integer 1 through N
RSC = $(\sec \theta)^{1/2}$
AN = θ = obliquity angle of chief ray in image space
EN = n = $w/2\pi$ = spatial frequency
X = wr(RSC) = argument of Bessel function
AJ = first order Bessel function, J_0
T = modulation transfer factor

*Contributed by Don Szeles, Institute of Science and Technology, The University of Michigan.

FLOW CHART – MODULATION TRANSFER FUNCTION



FORTRAN STATEMENTS—MODULATION TRANSFER FUNCTION

```
      DIMENSION R(21),F(21),AR(20),DE(20)
      READ1,N
    1 FORMAT (12)
      DO2,J=1,N
    2 READ3,E(J),R(J)
    3 FORMAT (E15.8,F15.8)
      READ4,RSC,AN
    4 FORMAT (E15.8,E15.8)
      NQ=N-1
      DO5,I=1,NQ
      AR(I)=(R(I+1)+R(I))/2.0
    5 DE(I)=E(I+1)-E(I)
    6 READ7,EN
    7 FORMAT (E15.8)
      SUM=0.0
      DO12,I=1,NQ
      X=6.2831853*EN*AR(I)*RSC
      IF(X-7.0)8,11,11
    8 Y=X*X/4.0
      W=Y
      EM=13.0
      SIG=-1.0
    9 Y=W*(Y/(EM*EM)+SIG)
      SIG=-SIG
      FM=EM-1.0
      IF(EM-1.0)10,10,9
   10 AJ=1.0-Y
      GO TO 12
   11 Y=1.0/(8.0*X)
      Z=Y*Y
      P=1.0-Z*(4.5-459.375*Z)
      Q=-Y*(1.0-Z*(37.5-7741.875*Z))
      SQ=SQRT(3.1415927*X)
      A=(P-Q)/SQ
      B=(P+Q)/SQ
      AJ=A*SIN(X)+B*COS(X)
   12 SUM=DF(I)*AJ+SUM
      T=SUM/F(N)
      PRINT13,EN,T,AN
   13 FORMAT (5HEN = ,E13.8,5H T = ,E13.8, 7H AN = ,E13.8)
      DO15,J=1,N
   15 PRINT16,J,E(J),J,R(J)
   16 FORMAT (2HE(,12,4H) = ,E13.8,6H R(,12,4H) = ,E13.8)
      GO TO 6
      END
```

45 CARDS

INPUT TO 1620 FREQUENCY RESPONSE

CARD NO.	FIELD 1 SIZE & TYPE	FIELD 1 VARIABLE	FIELD 2 SIZE & TYPE	FIELD 2 VARIABLE	FIELD 3 SIZE & TYPE	FIELD 3 VARIABLE
1	2, I*	N				
2 thru N+1	15, E*	E(J)	15, E	R(J)		
N + 2	15, E	RSC	15, E	AN		
N+3,N+4,...	15, E	EN				

Type I is an integer (e.g., 1, 12, 20).
Type E is a signed decimal number followed by a signed power of ten for relocating the decimal point (e.g., $-.842\ E-3 \rightarrow .000842$). A maximum of 8 places after the decimal point is allowed and the power of ten exponent must be > -99 and $< +99$.

*All numbers must be right justified in the field.

15.7. Optical Response Measuring Equipment

There is no standard way for measuring optical response. The techniques usually involve the preparation of a series of spatial sine waves and a method for measuring the image contrast. Two designs appear in the recent literature [26,27]; others can be developed. The essence of one device is quoted from [27]:

A stabilized tungsten ribbon filament lamp serves as a source, the filtered light passing through a condenser system uniformly illuminates a fine slit. The slit then serves as a self-luminous incoherent object. The lens under test forms an image of the slit in the front focal plane of a microscope objective. The microscope objective then presents the image to sinusoidally varying masks of fixed spatial frequency and amplitude which are mounted in frequency pairs on a rotating drum. The image transmitted by the sinusoids is integrated by a 931-A RCA photomultiplier tube. The diffuser in front of the phototube uniformly illuminates the photocathode elements. The output of the photomultiplier tube is then presented to a potentiometer pen recorder.

The area masks are arranged in frequency pairs, where one area mask is stepped 90° out of phase with the other. In this way ten spatial frequency pairs are arranged, all the peaks of the cosine area masks are in line with one another on the drum, then automatically the area masks representing the sine component would be in line. Thus when a slit image is presented to the area masks and its transmission measured and recorded, the peaks on the chart recording would then represent the real and imaginary parts of the transfer function for each spatial frequency available on the drum.

For an image which shows no phase shift effect, all the peaks on the chart recording corresponding to the transmission of the sinusoids (imaginary part) will be equal to each other and equal to the half amplitude zero spatial frequency area mask which in turn is equal to the average transmission of the area masks. These peaks represent the baseline for the measurement, and is the self-normalization constant for the measurement. The half-amplitude normalization target has an area transmittance equal to the average chart transmittance but is a zero frequency target. Thus at some frequency, when the image suffers a shift in phase a peak appears below the normalization curve; this phenomenon, most common when the system suffers from defect of focus, is called "spurious resolution." A curve connecting all the maxima of the cosines is then the real part of the transfer function of the system relative to the half-amplitude or average transmittance. A curve connecting all the sine terms is a plot of the imaginary part vs. frequency.

The drum carries ten frequency pairs, ranging from 0.201 lines/mm for the lowest to 8.20 lines/mm for the highest frequency. The apparatus is equipped with three

microscope objectives which magnify the frequencies such that it is possible to obtain a range from 1.79 lines/mm to 730 lines/mm. The drum rotates at two speeds. At the high speed the transfer function can be seen on an oscilloscope where rapid adjustments can be made for proper position and normalization. When the drum is switched to low speed, which is approximately 1 rpm, a permanent record can be made on the recorder. Thus it can be seen that after proper normalization, it takes only minutes to record directly the transfer function of an optical system. For an experienced operator a complete characteristic can be measured for a lens system for at least twenty positions through focus in one-half hour.

References

1. E. H. Linfoot and E. Wolf, *Proc. Phys. Soc.*, **66B,** 145 (1953).
2. T. Asakura and R. Barakat, *Japan J. Phys.*, **30,** 728.
3. R. Barakat, *Progress in Optics,* 1 North Holland Publishing Co., Amsterdam, 67-108. (1961) Chap. 3.
4. E. L. O'Neill, *J. Opt. Soc. Am.*, **46,** 285 (1956).
5. W. Steel, *Rev. optique,* **32,** 4(1953).
6. K. Miyamoto, *Progress in Optics,* 1 North Holland Publishing Co., Amsterdam, 33-66. (1961) Chap. 2.
7. L. J. Cutrona, E. N. Leith and L. J. Porcello, *Trans. I.R.E.*, **AC-4,** 137 (1959).
8. E. L. O'Neill, *Trans. I.R.E.*, **IT-2,** 56 (1956).
9. M. Born and E. Wolf, *Principles of Optics* Pergamon, New York, (1959).
10. H. H. Hopkins, *Wave Theory of Aberrations* Clarendon Press, Oxford, (1950).
11. A. Marechal and M. Francon, *Diffraction, Structure des Images* (Editions de la Revue d' Optique, Paris (1960).
12. R. Barakat, *J. Opt. Soc. Am.*, **52,** 985 (1962).
13. H. H. Hopkins, *Proc. Roy. Soc.*, **231,** 91 (1955).
14. G. Black and E. H. Linfoot, *Proc. Roy. Soc.*, **239A** 522 (1957).
15. R. Barakat, *J. Opt. Soc. Am.*, **53,** 324 (1963).
16. G. Parrent and C. Drane, *Optica Acta,* **3,** 195 (1956).
17. A. M. Goodbody, *Proc. Phys. Soc.*, **72,** 411 (1958).
18. M. De and B. K. Nath, *Optik,* **15,** 739 (1958).
19. R. Barakat and M. Morello, *J. Opt. Soc. Am.*, **52,** 1328A (1962).
20. R. Barakat, *J. Opt. Soc. Am.*, **51,** 152 (1961).
21. E. Wolf, *Phys. Soc. (London)*, Repts. Progr. in Phys., **14,** 95 (1951).
22. R. K. Luneberg, *Mathematical Theory of Optics,* Brown University Providence (1944).
23. R. Barakat, *J. Opt. Soc. Am.*, **52,** 276 (1962).
24. E. H. Linfoot, *J. Opt. Soc. Am.*, **46,** 740 (1957).
25. *Image Evaluation Techniques,* **III,** "Summer Course Notes," Institute of Optics, College of Engineering and Applied Science, The University of Rochester, Rochester, N.Y. (1963).
26. R. R. Shannon and A. H. Newman, *Appl. Opt.*, **2,** 365 (1963).
27. T. Tsurata, *Appl. Opt.*, **2,** 371 (1963).

Bibliography

Barakat, R., *J. Opt. Soc. Am.*, **52,** 264 (1962).

Conrady, A. E., *Applied Optics and Optical Design,* 1 Dover (1957).

Duffieux, P., *L'Integrale de Fourier et ses Applications a l'Optique,* Besancon, Faculte des Sciences, (1946).

Felgett, P., and E. H. Linfoot, *Trans. Roy. Soc.*, **247,** 369 (1954).

Marechal, A., *Rev. optique,* **27,** 73 (1948).

O'Neill, E. L., "Selected Topics in Optics and Communication Theory," Boston University Physical Research Laboratory (1958).

Picht, J., *Optische Abbildung* Braunschweig Vieweg, (1931).

Steward, G. C., *The Symmetrical Optical System,* Cambridge University Press, (1928).

Chapter 16

SPATIAL FREQUENCY FILTERING

James Alward

The University of Michigan

CONTENTS

16.1. Introduction 646
16.2. Basic Mathematical Relationships 646
16.2.1. Fourier Transform 646
16.2.2. Properties of the Two-Dimensional Fourier Transform 646
16.2.3. Two-Dimensional Fourier Transforms in Polar Coordinates . . . 647
16.2.4. Dirac Delta Function 648
16.2.5. Products and Convolutions 648
16.2.6. Autocorrelation Functions and Wiener Spectra. 649
16.3. Analysis of Spatial Frequency Filtering 649
16.3.1. Scanning Aperture Space Filters 650
16.3.2. Fixed-Field Moving-Reticle Space Filters 650
16.3.3. Scanning-Field Moving-Reticle Space Filters. 652
16.3.4. Circular Sectored Reticles 653
16.4. Fourier Transforms of Common Space Filters 654
16.4.1. Rectangular Aperture 654
16.4.2. Circular Aperture 654
16.4.3. Infinite Parallel-Spoke Reticle 654
16.4.4. Infinite Checkerboard Reticle. 654
16.4.5. Parallel-Spoke Reticle Limited by Rectangular Aperture 654
16.4.6. Parallel-Spoke Reticle Limited by Circular Aperture 655
16.4.7. Checkerboard Reticle Limited by Rectangular Aperture 655
16.4.8. Circular Sectored Reticles 655

16. Spatial Frequency Filtering

16.1. Introduction

In this chapter, the basic mathematical relations employed in spatial frequency filtering are summarized, the usual approaches to spatial filter analysis are presented, and the commonly encountered space filter expressions are tabulated. Almost no derivations are given here. Reference [1] contains a reasonably thorough, intuitive introduction to most of the concepts involved in spatial filtering.

16.2. Basic Mathematical Relationships

16.2.1. Fourier Transform. The Fourier transform and inverse transform for a two-dimensional spatial pattern $s(x,y)$ are:

$$S(k_x, k_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(x, y)\ e^{-j2\pi(k_x x + k_y y)}\, dx\, dy \tag{16-1}$$

$$s(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(k_x, k_y)\, e^{j2\pi(k_x x + k_y y)}\, dk_x\, dk_y \tag{16-2}$$

Spatial frequencies in the x and y directions are represented by k_x and k_y, respectively. The arbitrary spatial pattern $s(x,y)$ may be considered to be real for all incoherent infrared systems. $S(k_x,k_y)$ is in general complex. $|S(k_x,k_y)|$ is its amplitude term; $e^{j2\pi\phi(k_x,k_y)}$ is its phase term. Fourier transform pairs are denoted by a double headed arrow: $s(x,y) \leftrightarrow S(k_x,k_y)$. The condition for the existence of Eq. (16-1) and (16-2) is that the following inequality holds:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |s(x, y)|^2\, dx\, dy < \infty \tag{16-3}$$

16.2.2. Properties of the Two-Dimensional Fourier Transform. Some useful properties of the two-dimensional Fourier transform, which may be easily derived from the definition, are given in this section.

16.2.2.1. *Space Scaling.* If $s(x,y)$ and $S(k_x,k_y)$ are Fourier transform pairs, then the following transform relationship exists:

$$s(ax, by) \leftrightarrow \frac{1}{|a||b|} S\left(\frac{k_x}{a}, \frac{k_y}{b}\right) \tag{16-4}$$

16.2.2.2. *Space Shifting.* If $s(x,y)$ is shifted by a constant in each direction, its amplitude spectrum is unchanged, but its phase spectrum in each direction is modified by a term linear with space frequency.

$$s(x - x_0, y - y_0) \leftrightarrow S(k_x, k_y)\, e^{-j2\pi(k_x x_0 + k_y y_0)} \tag{16-5}$$

16.2.2.3. *Space-Frequency Shifting.* The corresponding transform pair for a shift in space frequency is:

$$e^{j2\pi(k_{x0}x + k_{y0}y)} s(x, y) \leftrightarrow S(k_x - k_{x0}, k_y - k_{y0}) \tag{16-6}$$

16.2.2.4. *Space and Space-Frequency Differentiation.*

$$\frac{d^m}{dx^m}\frac{d^n}{dy^n}s(x,y) \leftrightarrow (j2\pi k_x)^m\,(j2\pi k_y)^n\,S(k_x,k_y) \tag{16-7}$$

$$(-j2\pi x)^m(-j2\pi y)^n s(x,y) \leftrightarrow \frac{d^m}{dk_x{}^m}\frac{d^n}{dk_x{}^n}S(k_x,k_y) \tag{16-8}$$

16.2.2.5. *Conjugate Functions.* If $s(x,y)$ is a complex spatial function, then for its conjugate $s^*(x,y)$, the following transform pair holds:

$$s^*(x,y) \leftrightarrow S^*(-k_x,-k_y) \tag{16-9}$$

if $s(x,y)$ is real, then

$$s^*(x,y) = s(x,y) \leftrightarrow S(k_x,k_y) \tag{16-10}$$

$$S(k_x,k_y) = S^*(-k_x,-k_y) \tag{16-11}$$

$$S^*(k_x,k_y) = S(-k_x,-k_y) \tag{16-12}$$

16.2.2.6. *Symmetrical Spatial Functions.* If $s(x,y)$ is both real and symmetrical about the origin; *i.e.*, if $s(x,y) = s(-x,-y)$, then $S(k_x,k_y)$ is real and symmetrical, and

$$S(k_x,k_y) = 4\int_0^\infty\int_0^\infty s(x,y)\cos 2\pi(k_xx + k_yy)\,dx\,dy \tag{16-13}$$

$$s(x,y) = 4\int_0^\infty\int_0^\infty S(k_x,k_y)\cos 2\pi(k_xx + k_yy)\,dk_x\,dk_y \tag{16-14}$$

16.2.2.7. *Parseval's Theorem.* The two-dimensional expression for Parseval's theorem for two real spatial functions $s_1(x,y)$ and $s_2(x,y)$ is

$$\int_{-\infty}^\infty\int_{-\infty}^\infty s_1(x,y)s_2(x,y)\,dx\,dy = \int_{-\infty}^\infty\int_{-\infty}^\infty S_1(-k_x,-k_y)S_2(k_x,k_y)\,dk_x\,dk_y \tag{16-15}$$

$$= \int_{-\infty}^\infty\int_{-\infty}^\infty S_1(k_x,k_y)S_2(-k_x,-k_y)\,dk_x\,dk_y \tag{16-16}$$

$$= \int_{-\infty}^\infty\int_{-\infty}^\infty S_1{}^*(k_x,k_y)S_2(k_x,k_y)\,dk_x\,dk_y \tag{16-17}$$

$$= \int_{-\infty}^\infty\int_{-\infty}^\infty S_1(k_x,k_y)S_2{}^*(k_x,k_y)\,dk_x\,dk_y \tag{16-18}$$

16.2.3. Two-Dimensional Fourier Transforms in Polar Coordinates. The relations between rectangular coordinates (x,y) and (k_x,k_y) and polar coordinates (ρ,θ) and (k_ρ,ψ) are:

$$\left.\begin{aligned} x &= \rho\cos\theta & k_x &= k_\rho\cos\psi \\ y &= \rho\sin\theta & k_y &= k_\rho\sin\psi \\ x^2+y^2 &= \rho^2 & k_x{}^2+k_y{}^2 &= k_\rho{}^2 \\ dx\,dy &= \rho\,d\rho\,d\theta & dk_x\,dk_y &= k_\rho\,dk_\rho\,d\psi \end{aligned}\right\} \tag{16-19}$$

The Fourier transform has the form

$$S(k_\rho,\psi) = \int_0^\infty \int_0^{2\pi} s(\rho,\theta)\, e^{-j2\pi k_\rho \rho(\cos\theta\cos\psi + \sin\theta\sin\psi)} \rho\, d\theta\, d\rho$$

$$= \int_0^\infty \int_0^{2\pi} s(\rho,\theta)\, e^{-j2\pi k_\rho \rho\cos(\theta-\psi)} \rho\, d\theta\, d\rho \qquad (16\text{-}20)$$

The inverse Fourier transform has the form

$$s(\rho,\theta) = \int_0^\infty \int_0^{2\pi} S(k_\rho,\psi)\, e^{j2\pi k_\rho \rho\cos(\theta-\psi)} k_\rho\, d\psi\, dk_\rho \qquad (16\text{-}21)$$

If the spatial function is not dependent on θ, then the Fourier transform pair for $s(\rho)$ takes the form

$$S(k_\rho) = \int_0^\infty s(\rho) J_0(k_\rho\rho)\rho\, d\rho \qquad (16\text{-}22)$$

$$s(\rho) = \int_0^\infty S(k_\rho) J_0(k_\rho\rho) k_\rho\, dk_\rho \qquad (16\text{-}23)$$

where J_0 is the Bessel function of the first kind and order zero.

16.2.4. Dirac Delta Function. The spatial Dirac delta function $\delta(x,y)$ is used to represent a finite energy source or a finite transmittance concentrated into an arbitrarily small region of the plane. For example, $\delta(x-x_0,y-y_0)$ is used to represent a "point" source at (x_0, y_0). The energy density at (x_0,y_0) is infinitely large.

Mathematically, the delta function can be defined in terms of its sifting property:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(x-x_0, y-y_0) s(x,y)\, dx\, dy = s(x_0, y_0) \qquad (16\text{-}24)$$

where $s(x,y)$ is an arbitrary function continuous at (x_0,y_0). In terms of this definition, the Fourier transform of $\delta(x-x_0,y-y_0)$ is easily found to be $e^{-j2\pi(k_x x_0 + k_y y_0)}$ thus establishing a Fourier transform pair for the Dirac delta function.

A delta function in terms of only one spatial variable $\delta(x - x_0)$ may be interpreted physically as a line source or a line transmittance of infinite length and arbitrarily small width, along the line $x = x_0$.

Another property of the delta function is:

$$\delta(ax,by) = \frac{1}{|a||b|}\,\delta(x,y) \qquad (16\text{-}25)$$

16.2.5. Products and Convolutions. If the inverse Fourier transform of the product of two space-frequency functions $S_1(k_x,k_y)$ and $S_2(k_x,k_y)$ is taken, the result is the convolution of the inverse Fourier transforms of the two functions:

$$s(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} s_1(\xi,\eta) s_2(x-\xi, y-\eta)\, d\xi\, d\eta \qquad (16\text{-}26)$$

Similarly, if the Fourier transform of the product of two spatial functions $s_1(x,y)$ and $s_2(x,y)$ is taken, the result is the convolution of the Fourier transforms of the two functions:

$$S(k_x, k_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S_1(k_\xi, k_\eta) S_2(k_x - k_\xi, k_y - k_\eta)\, dk_\xi\, dk_\eta \tag{16-27}$$

16.2.6. Autocorrelation Functions and Wiener Spectra. The autocorrelation function $w(\xi,\eta)$ is defined as a joint mean of a random process [1].

$$w(\xi, \eta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s_1 s_2 p_2(s_1, \xi; s_2, \eta)\, ds_1\, ds_2 \tag{16-28}$$

where s_1 and s_2 are two sample functions evaluated at two randomly chosen points, (x_1,y_1) and (x_2,y_2) respectively, and p_2 is the second-order joint probability density function of the random process which generated s_1 and s_2. ξ and η are equal to $x_1 - x_2$ and $y_1 - y_2$, the displacement coordinates between (x_1,y_1) and (x_2,y_2), respectively. The random process is assumed to be stationary in terms of second-order statistics. If the random process is also ergodic, the autocorrelation function may be represented in terms of an average over space of a single sample function of the random process:

$$w(\xi, \eta) = \lim_{\substack{A \to \infty \\ B \to \infty}} \frac{1}{4AB} \int_{-A}^{A} \int_{-B}^{B} s(x,y) s(x+\xi, y+\eta)\, dx\, dy \tag{16-29}$$

The condition for the existence of Eq. (16-28) and (16-29) is that total mean square average of the random function be bounded; *i.e.,* that

$$\lim_{\substack{A \to \infty \\ B \to \infty}} \frac{1}{4AB} \int_{-A}^{A} \int_{-B}^{B} |s(x,y)|^2\, dx\, dy < \infty \tag{16-30}$$

which is a less stringent condition than expression (16-3). An autocorrelation function for functions satisfying (16-3) may be represented as follows:

$$w(\xi, \eta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(x, y) s(x+\xi, y+\eta)\, dx\, dy \tag{16-31}$$

The Wiener spectrum, $W(k_x,k_y)$, the analogy of the power spectrum in electrical systems analysis, is defined simply as the Fourier transform of the autocorrelation function:

$$W(k_x\, k_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w(\xi, \eta) e^{-j2\pi(k_x \xi + k_y \eta)}\, d\xi\, d\eta \tag{16-32}$$

If the spatial pattern $s(x,y)$ satisfies expression (16-3) the Wiener spectrum may be represented directly in terms of $S(k_x,k_y)$:

$$W(k_x, k_y) = |S(k_x, k_y)|^2 \tag{16-33}$$

Since the Fourier transform of a random (continuing) function does not exist [1], the Wiener spectrum must be used whenever a space frequency representation is needed.

16.3. Analysis of Spatial Frequency Filtering

The basic expression for the time-varying output $v(t)$ of a spatio-temporal filter is

$$v(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} r(x, y, t) s(x, y)\, dx\, dy \tag{16-34}$$

where $r(x,y,t)$ is a general expression for a spatio-temporal filter, and $s(x,y)$ is an arbitrary input scene. The infinite limits are for generality; actually $r(x,y,t)$ is zero

outside a finite region. This expression applies to scanning-aperture space filters, to fixed-field moving-reticle space filters, and to scanning-field moving-reticle space filters.

16.3.1. Scanning-Aperture Space Filters. For a scanning aperture moving with a velocity u_x parallel to the x axis, $r(x,y,t)$ is represented as $r(x-u_x t,y)$ and the output is given by

$$v(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} r(x-u_x t, y)s(x, y)\, dx\, dy \tag{16-35}$$

The Fourier transform $V(f)$ (also called the output spectrum) of this expression is given by

$$V(f) = \frac{1}{u_x} \int_{-\infty}^{\infty} R^*\left(\frac{f}{u_x}, k_y\right) S\left(\frac{f}{u_x}, k_y\right) dk_y \tag{16-36}$$

If the input scene is represented only in terms of its Wiener spectrum $W(k_x,k_y)$, the output power spectrum $\Phi(f)$ is

$$\Phi(f) = \frac{1}{u_x} \int_{-\infty}^{\infty} \left|R\left(\frac{f}{u_x}, k_y\right)\right|^2 W\left(\frac{f}{u_x}, k_y\right) dk_y \tag{16-37}$$

If there is also a velocity component u_y in the y direction, the expressions for $V(f)$ and $\Phi(f)$ are

$$V(f) = \frac{1}{u_x u_y} R^*\left(\frac{f}{u_x}, \frac{f}{u_y}\right) S\left(\frac{f}{u_x}, \frac{f}{u_y}\right) \tag{16-38}$$

$$\Phi(f) = \frac{1}{u_x u_y}\left|R\left(\frac{f}{u_x}, \frac{f}{u_y}\right)\right|^2 W\left(\frac{f}{u_x}, \frac{f}{u_y}\right) \tag{16-39}$$

16.3.2. Fixed-Field Moving-Reticle Space Filters. One type of fixed-field moving-reticle space filter is simply an infinite reticle pattern $r_\infty(x-u_x t,y)$ scanning over the limited input scene $s(x,y)r_a(x,y)$, where $r_a(x,y)$ is the fixed field of view of the space filter.

The Fourier transform of an infinite parallel-spoke square-wave reticle, assuming the spokes are parallel to the y axis, is an array of delta functions (along the k_x axis) whose magnitudes are determined from the Fourier series of a square wave:

$$R_\infty(k_x, k_y) = \frac{1}{2}\,\delta(k_x, k_y) + \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)\pi}\,\delta(k_x - [2n-1]k_0, k_y) + \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)\pi}\,\delta(k_x + [2n-1]k_0, k_y) \tag{16-40}$$

where k_0 is the fundamental space frequency, the period being the width of one spoke pair. The output spectrum $V(f)$ from scanning the endless reticle over the limited scene with a velocity u_x in the x direction is:

$$V(f) = \frac{1}{u_x}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(\zeta_x, \zeta_y)R_a\left(\frac{f}{u_x} - \zeta_x, -\zeta_y\right) d\zeta_x\, d\zeta_y\right] \times \left\{\frac{1}{2}\,\delta\left(\frac{f}{u_x}\right) + \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)\pi}\,\delta\left[\frac{f}{u_x} - (2n-1)k_0\right] + \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)\pi}\,\delta\left[\frac{f}{u_x} + (2n-1)k_0\right]\right\} \tag{16-41}$$

The output is periodic, hence the expression is simply the transform of a Fourier series whose coefficients are determined from the square wave, but are further modified by the spectrum of the limited scene. If the input scene is represented only in the form of a Wiener spectrum $W(k_x,k_y)$, the output power spectrum $\Phi(f)$ is

$$\Phi(f) = \frac{1}{u_x}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(\zeta_x,\zeta_y)\left|R_a\left(\frac{f}{u_x}-\zeta_x,-\zeta_y\right)\right|^2 d\zeta_x\, d\zeta_y\right]$$

$$\times\left\{\frac{1}{4}\,\delta\left(\frac{f}{u_x}\right)+\sum_{n=1}^{\infty}\frac{\delta\left[\frac{f}{u_x}-(2n-1)k_0\right]}{[(2n-1)\pi]^2}+\sum_{n=1}^{\infty}\frac{\delta\left[\frac{f}{u_x}+(2n-1)k_0\right]}{[(2n-1)\pi]^2}\right\} \tag{16-42}$$

The Fourier transform of an infinite checkerboard square-wave reticle, assuming that the pattern is oriented so the square edges are parallel to the coordinate axes, is a double array of delta functions along the 45° lines which bisect the right angles formed by the coordinates in the (k_x,k_y) plane. The magnitudes are determined from the Fourier series of a symmetrical triangle wave [1].

$$R_\infty(k_x,k_y) = \frac{1}{2}\,\delta(k_x,k_y)+\sum_{n=1}^{\infty}\frac{1}{[(2n-1)\pi]^2}\,\delta[k_x-(2n-1)k_0,\,k_y-[2n-1]k_0$$

$$+\sum_{n=1}^{\infty}\frac{1}{[(2n-1)\pi]^2}\,\delta[k_x+(2n-1)k_0,\,k_y+(2n-1)k_0]$$

$$+\sum_{n=1}^{\infty}\frac{1}{[(2n-1)\pi]^2}\,\delta[k_x-(2n-1)k_0,\,k_y+(2n-1)k_0]$$

$$+\sum_{n=1}^{\infty}\frac{1}{[(2n-1)\pi]^2}\,\delta[k_x+(2n-1)k_0,\,k_y-(2n-1)k_0] \tag{16-43}$$

where k_0 is the fundamental space frequency in either the x or y directions, the period being twice the width of one checkerboard square. The output spectrum $V(f)$ from scanning the endless checkerboard with a velocity of u_x in the x direction over a scene limited by $r_a\,(x,y)$ is:

$$V(f) = \frac{1}{u_x}\frac{1}{2}\,\delta\left(\frac{f}{u_x}\right)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(\zeta_x,\zeta_y)R_a\left(\frac{f}{u_x}-\zeta_x,-\zeta_y\right)d\zeta_x\,d\zeta_y$$

$$+\frac{1}{u_x}\sum_{n=1}^{\infty}\frac{1}{[(2n-1)\pi]^2}\,\delta\left[\frac{f}{u_x}-(2n-1)k_0\right]\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(\zeta_x,\zeta_y)$$

$$R_a\left[\frac{f}{u_x}-\zeta_x,(2n-1)k_0-\zeta_y\right]d\zeta_x\,d\zeta_y$$

$$+\frac{1}{u_x}\sum_{n=1}^{\infty}\frac{1}{[(2n-1)\pi]^2}\,\delta\left[\frac{f}{u_x}+(2n-1)k_0\right]\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(\zeta_x,\zeta_y)$$

$$R_a\left[\frac{f}{u_x}-\zeta_x,-(2n-1)k_0-\zeta_y\right]d\zeta_x\,d\zeta_y$$

$$+\frac{1}{u_x}\sum_{n=1}^{\infty}\frac{1}{[(2n-1)\pi]^2}\,\delta\left[\frac{f}{u_x}-(2n-1)k_0\right]\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}S(\zeta_x,\zeta_y)$$

$$R_a\left[\frac{f}{u_x}-\zeta_x,-(2n-1)k_0-\zeta_y\right]d\zeta_x\,d\zeta_y$$

$$+\frac{1}{u_x}\sum_{n=1}^{\infty}\frac{1}{[(2n-1)\pi]^2}\,\delta\left[\frac{f}{u_x}+(2n-1)k_0\right]\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}S(\zeta_x,\zeta_y)$$

$$R_a\left[\frac{f}{u_x}-\zeta_x,(2n-1)k_0-\zeta_y\right]d\zeta_x\,d\zeta_y \qquad (16\text{-}44)$$

If the input is described only in terms of its Wiener spectrum, the output power spectrum $\Phi(f)$ corresponding to Eq. (16-44) is

$$\Phi(f)=\frac{1}{u_x}\frac{1}{4}\,\delta\left(\frac{f}{u_x}\right)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}W(\zeta_x,\zeta_y)\left|R_a\left(\frac{f}{u_x}-\zeta_x,-\zeta_y\right)\right|^2 d\zeta_x\,d\zeta_y$$

$$+\frac{1}{u_x}\sum_{n=1}^{\infty}\frac{1}{[(2n-1)\pi]^4}\,\delta\left[\frac{f}{u_x}-(2n-1)k_0\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}W(\zeta_x,\zeta_y)\right.$$

$$\left|R_a\left[\frac{f}{u_x}-\zeta_x,(2n-1)k_0-\zeta_y\right]\right|^2 d\zeta_x\,d\zeta_y$$

$$+\frac{1}{u_x}\sum_{n=1}^{\infty}\frac{1}{[(2n-1)\pi]^4}\,\delta\left[\frac{f}{u_x}+(2n-1)k_0\right]\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}W(\zeta_x,\zeta_y)$$

$$\left|R_a\left[\frac{f}{u_x}-\zeta_x,-(2n-1)k_0-\zeta_y\right]\right|^2 d\zeta_x\,d\zeta_y$$

$$+\frac{1}{u_x}\sum_{n=1}^{\infty}\frac{1}{[(2n-1)\pi]^4}\,\delta\left[\frac{f}{u_x}-(2n-1)k_0\right]\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}W(\zeta_x,\zeta_y)$$

$$\left|R_a\left[\frac{f}{u_x}-\zeta_x,-(2n-1)k_0-\zeta_y\right]\right|^2 d\zeta_x\,d\zeta_y$$

$$+\frac{1}{u_x}\sum_{n=1}^{\infty}\frac{1}{[(2n-1)\pi]^4}\,\delta\left[\frac{f}{u_x}+(2n-1)k_0\right]\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}W(\zeta_x,\zeta_y)$$

$$\left|R_a\left[\frac{f}{u_x}-\zeta_x,(2n-1)k_0-\zeta_y\right]\right|^2 d\zeta_x\,d\zeta_y \qquad (16\text{-}45)$$

16.3.3. Scanning-Field Moving-Reticle Space Filters. The scanning-field moving-reticle space filter has both a moving reticle and a scanning aperture; however, the velocities of the two elements are not necessarily equal nor even in the same direction. In the previous case of the reticle scanning a scene limited by a fixed aperture, the time-varying output was periodic. The effect of having the aperture move is one of modulating the periodic signal and thus spreading the signal energy into frequency bands about the original signal harmonics [2]. Two scanning situations are considered below.

The first situation concerns both the reticle $r(x,y)$ and the aperture field stop $a(x,y)$ moving in the x direction. The velocity of the reticle with respect to the scene is u_{xr},

the velocity of the aperture with respect to the scene is u_{xa}, and the velocity of the reticle with respect to the aperture is u_{xra}, where $u_{xra} = u_{xr} - u_{xa}$. Usually u_{xr} is greater than u_{xa}. The time-varying output is given by

$$v(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(x,y)\,a(x-u_{xa}t, y)\,r(x-u_{xr}t, y)\,dx\,dy \tag{16-46}$$

The corresponding output spectrum $V(f)$ is

$$V(f) = \frac{1}{u_{xra}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(k_x, k_y) A^*\left(\frac{k_x u_{xr} - f}{u_{xra}}, k_y - k'_y\right) R^*\left(\frac{f - k_x u_{xa}}{u_{xa}}, k'_y\right) dk_x\,dk_y\,dk'_y \tag{16-47}$$

The second situation is concerned with the reticle scanning in the y direction with a velocity of u_{yr} with respect to the scene, and the aperture moving in the x direction with a velocity of u_{xa} with respect to the scene. Also, the reticle does not move with respect to the aperture in the x direction, so that the reticle also has a velocity of u_{xa} in the x direction with respect to the *scene*. The time-varying output is given by

$$v(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(x,y)\,a(x-u_{xa}t, y)\,r(x-u_{xa}t, y-u_{yr}t)\,dx\,dy \tag{16-48}$$

The corresponding output spectrum is

$$V(f) = \frac{1}{u_{yr}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(k_x, k_y) A^*\left(k_x - k'_x, k_y + \frac{k_x u_{xa} - f}{u_{yr}}\right)$$

$$R^*\left(k'_x, \frac{f - k_x u_{xa}}{u_{yr}}\right) dk'_x\,dk_y\,dk_x \tag{16-49}$$

Expressions (16-47) and (16-49) involve a double convolution in the k_y and the k_x directions, respectively.

For inputs represented only in terms of their Wiener spectra, the output power spectrum expressions corresponding to Eq. (16-47) and (16-49) are, respectively:

$$\Phi(f) = \frac{1}{u_{xra}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(k_x, k_y) \left|A\left(\frac{k_x u_{xr} - f}{u_{xra}}, k_y - k'_y\right)\right|^2$$

$$\left|R\left(\frac{f - k_x u_{xa}}{u_{xa}}, k'_y\right)\right|^2 dk_x\,dk_y\,dk'_y \tag{16-50}$$

$$\Phi(f) = \frac{1}{u_{yr}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(k_x, k_y) \left|A\left(k_x - k'_x, k_y + \frac{k_x u_{xa} - f}{u_{yr}}\right)\right|^2$$

$$\left|R\left(k'_x, \frac{f - k_x u_{xa}}{u_{yr}}\right)\right|^2 dk_x{}^1\,dk_y\,dk_x \tag{16-51}$$

16.3.4. Circular Sectored Reticles. The circular sectored reticle, which is also known as the "wagon-wheel" reticle or the episcotister, is shown in Section 16.4. This reticle is best represented in polar coordinates as a periodic variation in the θ direction. The Fourier transform $R(k_\rho, \psi)$ of a circular sectored reticle with n black-white spoke pairs and with a radius of a is given by

$$R(k_\rho,\psi) = \frac{aJ_1(2\pi ak_\rho)}{2k_\rho} + \frac{1}{\pi^2 k_\rho^2} \sum_{p=1}^{\infty} \frac{(-j)^{(2p-1)n}}{2p-1} \sin (2p-1)n\psi \int_0^{2\pi ak_\rho} zJ_{(2p-1)n}(z)\, dz \tag{16-52}$$

$J_{(2p-1)n}(z)$ denotes the Bessel function of the first kind and order $(2p-1)n$. $R(k_\rho,\psi)$ is a complex function when n is odd, and is real when n is even. If the original reticle function is rotated with an angular velocity ωt so that $r(\rho,\theta)$ becomes $r(\rho,\theta-\omega t)$, then $R(k_\rho,\psi)$ becomes $R(k_\rho,\psi+\omega t)$. In other words, the Fourier transform rotates with the same velocity in the opposite direction.

16.4. Fourier Transforms of Common Space Filters

Ten pairs of the more common space filters and their Fourier transforms are illustrated and summarized below for use in the mathematical expressions of Section 16.3, especially Section 16.3.2. Solid black in the figures indicates a transmission of 0; white indicates a transmission of 1.

16.4.1. Rectangular Aperture. Figure 16-1(a) show a simple rectangular aperture with dimensions $a \times b$. The Fourier transform of this aperture is

$$R_a(k_x, k_y) = \frac{\sin \pi ak_x \sin \pi bk_y}{\pi^2 k_x k_y} \tag{16-53}$$

One quadrant of $|R_a(k_x, k_y)|$ is shown in Fig. 16-1(b).

16.4.2. Circular Aperture. Figure 16-2(a) shows a circular aperture of radius a. The Fourier transform of this aperture is

$$R_a(k_x, k_y) = \frac{aJ_1(2\pi a\sqrt{k_x^2 + k_y^2})}{\sqrt{k_x^2 + k_y^2}} \tag{16-54}$$

One quadrant of $|R_a(k_x, k_y)|$ is shown in Fig. 16-2(b).

16.4.3. Infinite Parallel-Spoke Reticle. A parallel-spoke reticle which is infinite in extent in both directions is shown in Fig. 16-3(a). Its Fourier transform is given by Eq. (16-40) and is shown in Fig. 16-3(b).

16.4.4. Infinite Checkerboard Reticle. A checkerboard reticle which is infinite in extent in both directions is shown in Fig. 16-4(a). Its Fourier transform is given by Eq. 16-43 and is shown in Fig. 16-4(b).

16.4.5. Parallel-Spoke Reticle Limited by Rectangular Aperture. A parallel-spoke reticle limited by a rectangular aperture of dimensions $a \times b$ is shown in Fig. 16-5(a). The Fourier transform of this reticle is given by either of the following expressions.

$$R(k_x, k_y) = \frac{1}{2} \frac{\sin \pi ak_x}{\pi k_x} \frac{\sin \pi bk_y}{\pi k_y} + \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)\pi} \frac{\sin \pi a[k_x - (2n-1)k_0]}{\pi[k_x - (2n-1)k_0]} \frac{\sin \pi bk_y}{\pi k_y}$$

$$+ \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)\pi} \frac{\sin \pi a[k_x + (2n-1)k_0]}{\pi[k_x + (2n-1)k_0]} \frac{\sin \pi bk_y}{\pi k_y} \tag{16-55}$$

$$R(k_x, k_y) = \frac{\sin \pi bk_y}{\pi k_y} \frac{\sin \pi(x_0/2)k_x}{\pi k_x} \frac{\sin \pi ak_x}{\sin \pi x_0 k_x} \tag{16-56}$$

Figure 16-5(b) shows $|R(k_x, k_y)|$ for $k_x > 0$.

16.4.6. Parallel-Spoke Reticle Limited by Circular Aperture. A parallel-spoke reticle limited by a circular aperture of radius a is shown in Fig. 16-6(a). The Fourier transform of this reticle is

$$R(k_x, k_y) = \frac{a}{2}\frac{J_1(2\pi a\sqrt{k_x^2 + k_y^2})}{\sqrt{k_x^2 + k_y^2}} + \sum_{n=1}^{\infty}\frac{(-1)^{n-1}a}{(2n-1)\pi}\frac{J_1(2\pi a\sqrt{[k_x - (2n-1)k_0]^2 + k_y^2}}{\sqrt{[k_x - (2n-1)k_0]^2 + k_y^2}}$$

$$+ \sum_{n=1}^{\infty}\frac{(-1)^{n-1}a}{(2n-1)\pi}\frac{J_1(2\pi a\sqrt{[k_x + (2n-1)k_0]^2 + k_y^2}}{\sqrt{[k_x + (2n-1)k_0]^2 + k_y^2}} \tag{16-57}$$

One quadrant of $|R(k_x,k_y)|$ is shown in Fig. 16-6(b).

16.4.7. Checkerboard Reticle Limited by Rectangular Aperture. A checkerboard reticle limited by a rectangular aperture with dimensions $a \times b$ is shown in Fig. 16-7(a). The Fourier transform of this reticle is given by either of the following two expressions:

$$R(k_x, k_y) = \frac{1}{2}\frac{\sin \pi a k_x}{\pi k_x}\frac{\sin \pi b k_y}{\pi k_y}$$

$$+ \sum_{n=1}^{\infty}\frac{1}{[(2n-1)\pi]^2}\frac{\sin \pi a[k_x - (2n-1)k_0]}{\pi[k_y - (2n-1)k_0]}\frac{\sin \pi b[k_y - (2n-1)k_0]}{\pi[k_y - (2n-1)k_0]}$$

$$+ \sum_{n=1}^{\infty}\frac{1}{[(2n-1)\pi]^2}\frac{\sin \pi a[k_x + (2n-1)k_0]}{\pi[k_x + (2n-1)k_0]}\frac{\sin \pi b[k_y + (2n-1)k_0]}{\pi[k_y + (2n-1)k_0]}$$

$$+ \sum_{n=1}^{\infty}\frac{1}{[(2n-1)\pi]^2}\frac{\sin \pi a[k_x - 2n-1)k_0]}{\pi[k_x - (2n-1)k_0]}\frac{\sin \pi b[k_y + (2n-1)k_0]}{\pi[k_y + (2n-1)k_0]}$$

$$+ \sum_{n=1}^{\infty}\frac{1}{[(2n-1)\pi]^2}\frac{\sin \pi a[k_x + (2n-1)k_0]}{\pi[k_x + (2n-1)k_0]}\frac{\sin \pi b[k_y - (2n-1)k_0]}{\pi[k_y - (2n-1)k_0]} \tag{16-58}$$

$$R(k_x, k_y) = 2\frac{\sin \pi(x_0/2)k_x}{\pi k_x}\frac{\sin \pi(x_0/2)k_y}{\pi k_y}\frac{\sin \pi a k_x}{\sin \pi x_0 k_x}\frac{\sin \pi b k_y}{\sin \pi x_0 k_y}\cos \pi\frac{x_0}{2}(k_x + k_y) \tag{16-59}$$

$|R(k_x,k_y)|$ is shown in Fig. 16-7(b) for $k_x > 0$.

16.4.8. Circular Sectored Reticles. Circular sectored reticles with one, two, and four spoke pairs are shown in Figs. 16-8(a), 16-9(a), and 16-10(a). The Fourier transform of a circular sectored reticle with n spoke pairs is given by Eq. (16-52), which consists of a term independent of n and a summation of terms dependent on n. The term which is independent of n is simply one-half of expression (16-54), which is shown in Fig. 16-2(b). Figures 16-8(b), 16-9(b), and 16-10(b) show only the summation terms of $R(k_\rho, \psi)$ after the first term has been removed.

References

1. *IRIA State-of-the-Art Report on Spatial Frequency Filtering,* Report No. 2389-87-T. The University of Michigan, Ann Arbor, Michigan (To be published).
2. V. J. Ashby, *The Principles of Space Filtering in the Image Plane of a Simple Optical System,* Space Technology Laboratories, Canoga Park, Calif. (1961).

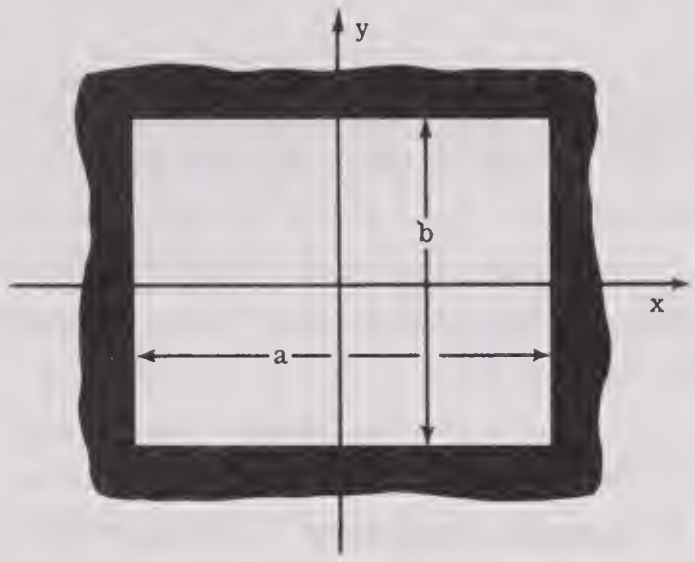


FIG. 16-1(*a*)

FIG. 16-1(*b*)

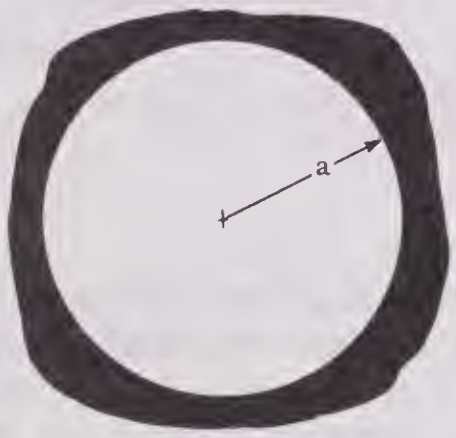


FIG. 16-2(*a*)

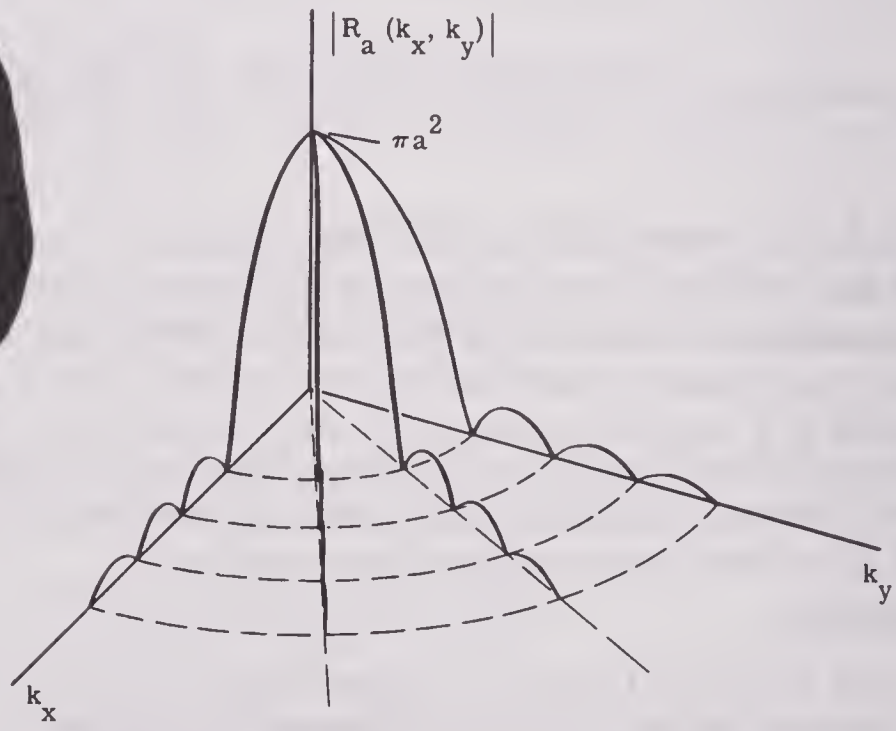


FIG. 16-2(*b*)

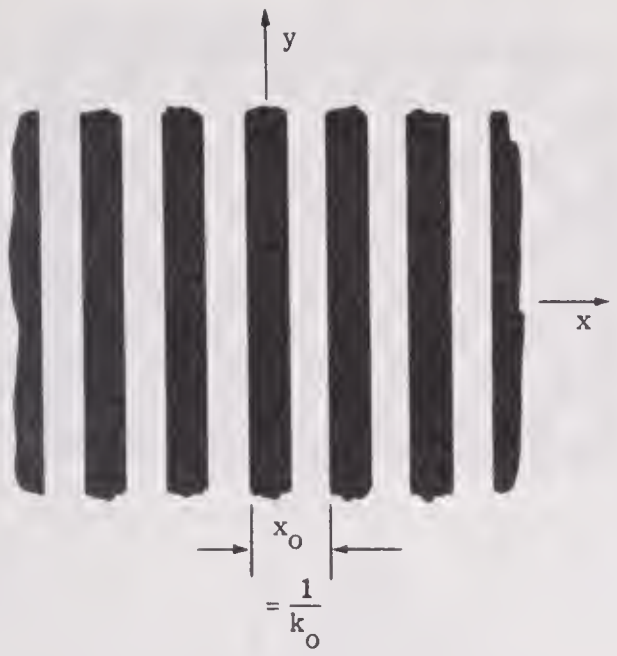


FIG. 16-3(*a*)

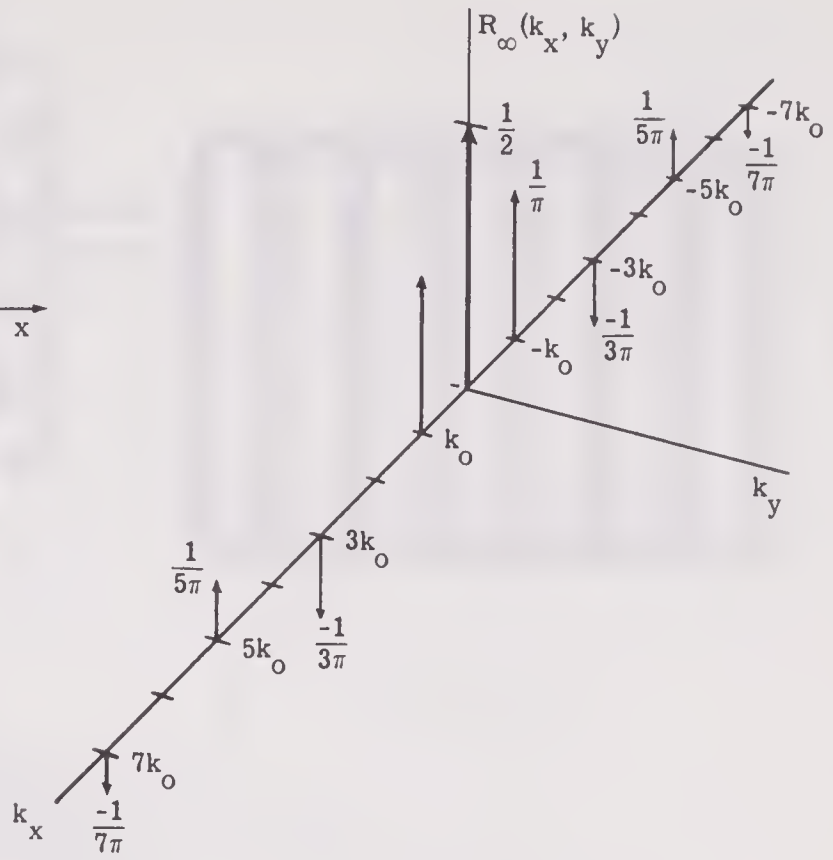


FIG. 16-3(*b*)

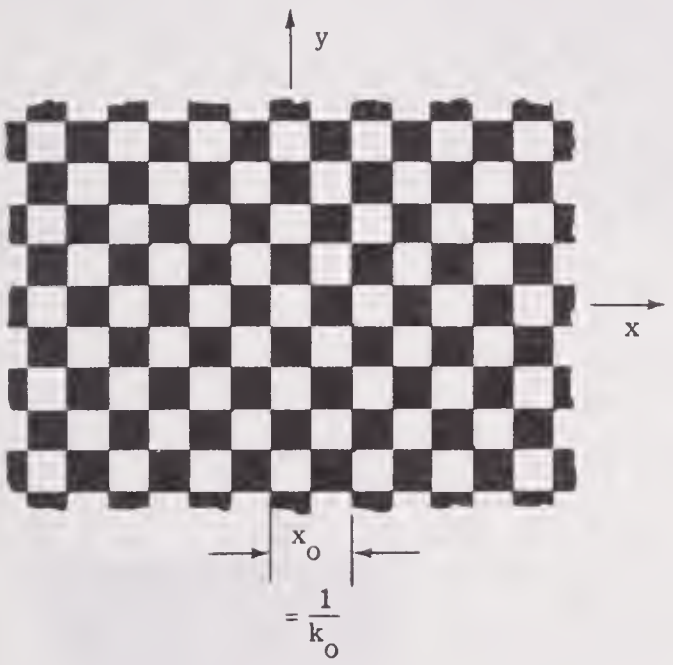


FIG. 16-4(*a*)

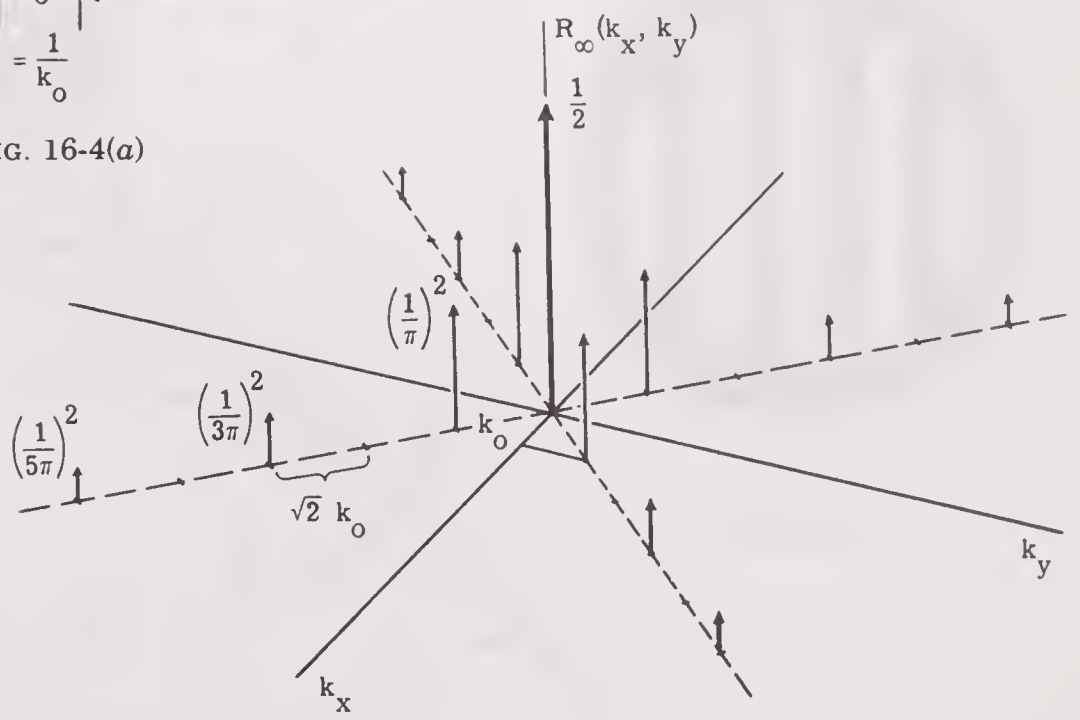


FIG. 16-4(*b*)

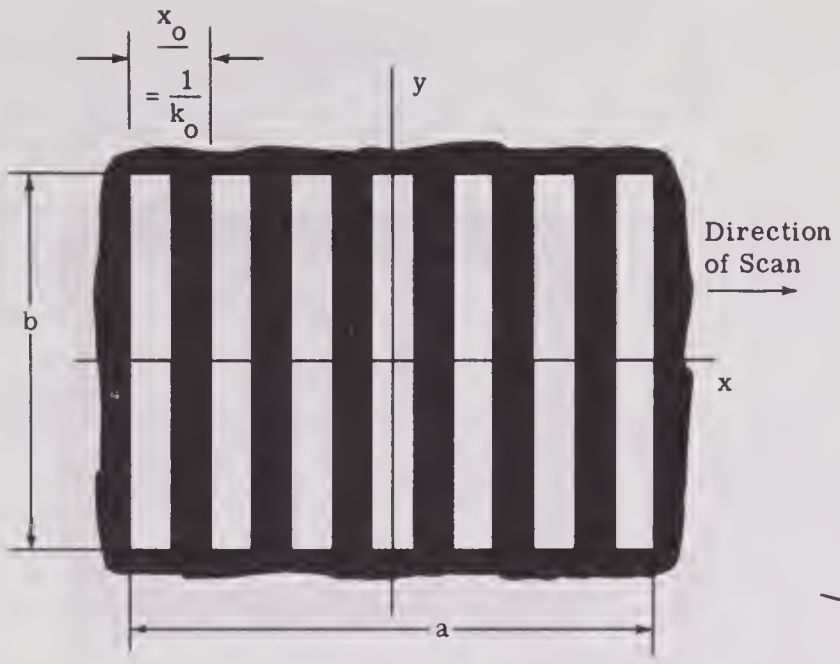


FIG. 16-5(*a*)

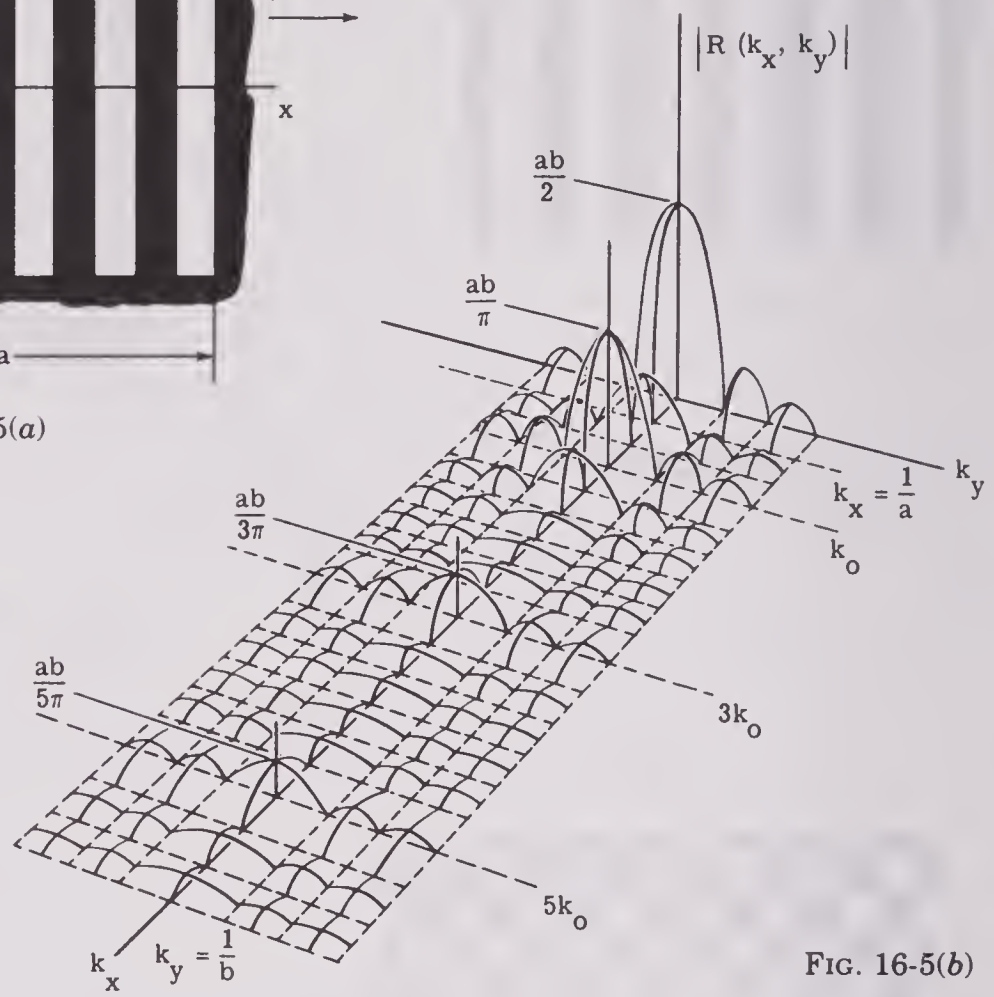


FIG. 16-5(*b*)

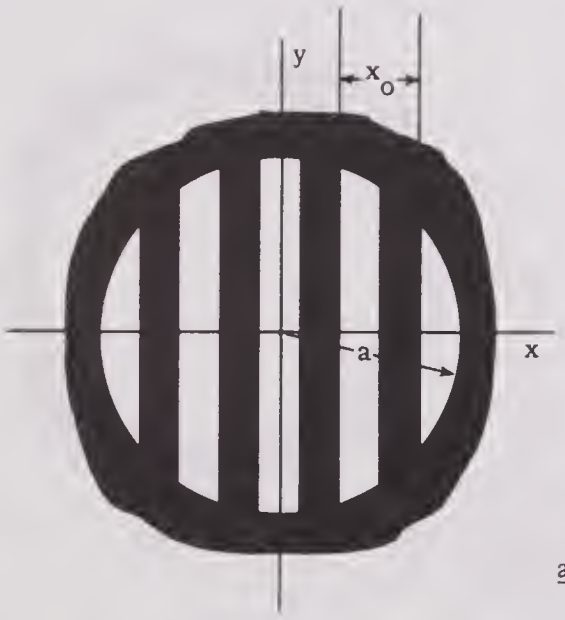


FIG. 16-6(*a*)

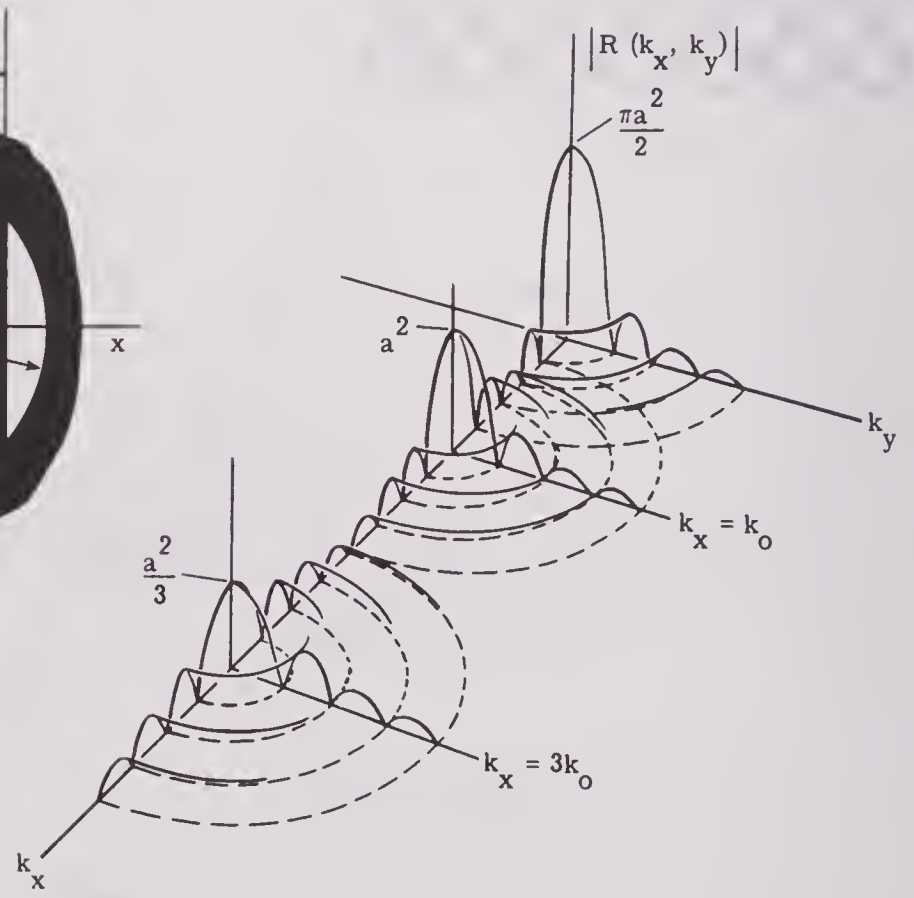


FIG. 16-6(*b*)

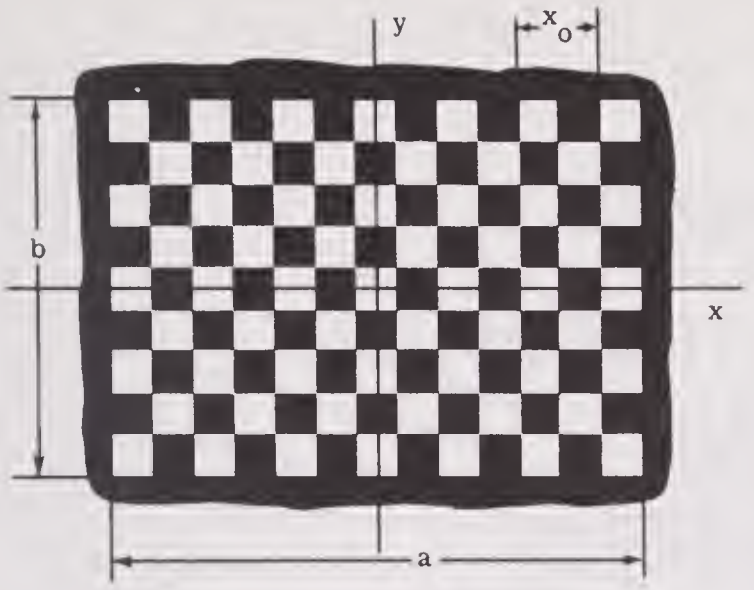


FIG. 16-7(*a*)

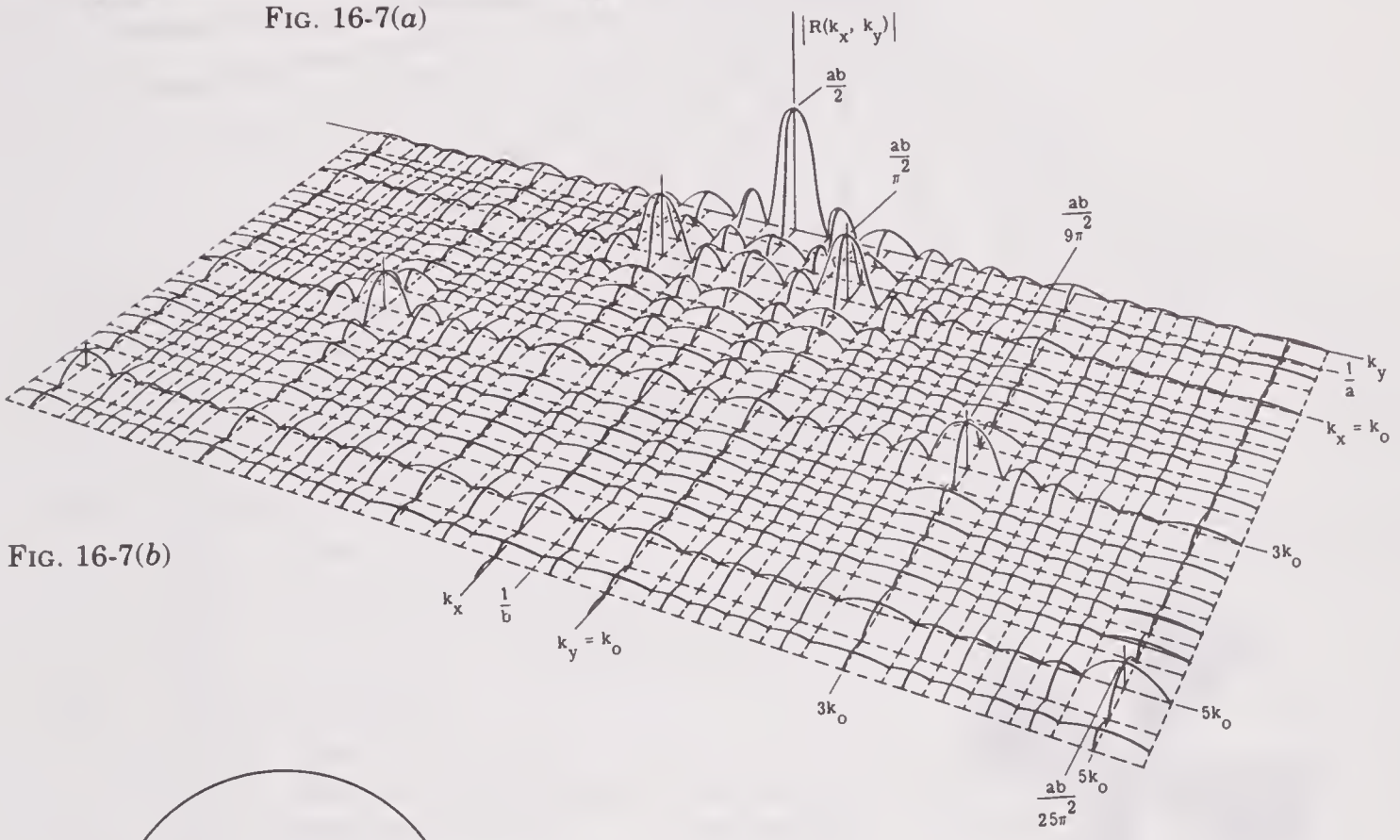


FIG. 16-7(*b*)

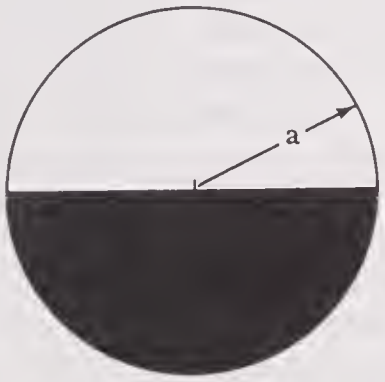


FIG. 16-8(*a*)

FIG. 16-8(*b*)

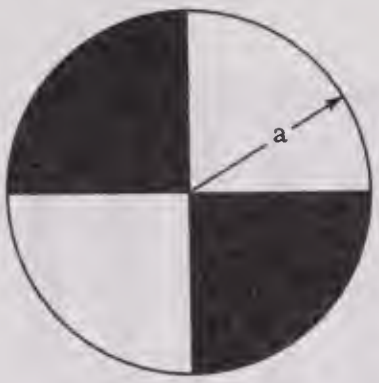


FIG. 16-9(*a*)

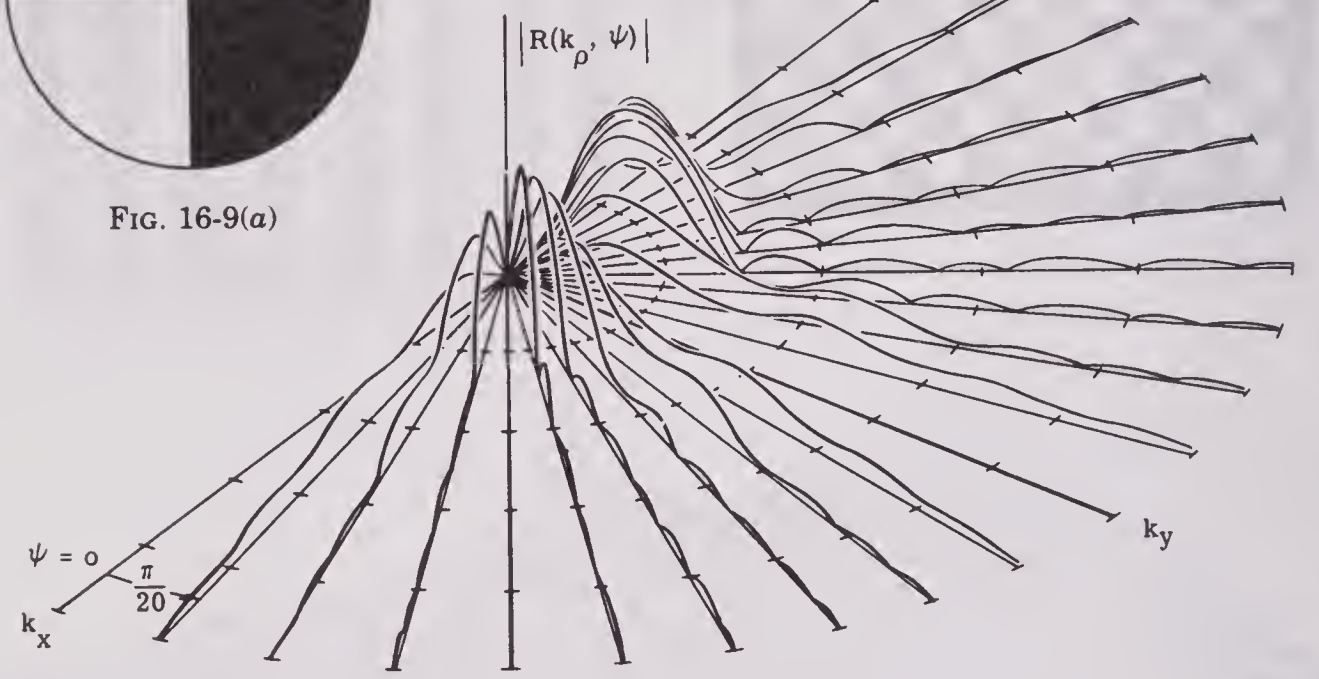


FIG. 16-9(*b*)

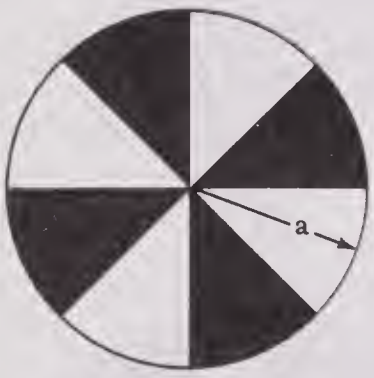


FIG. 16-10(*a*)

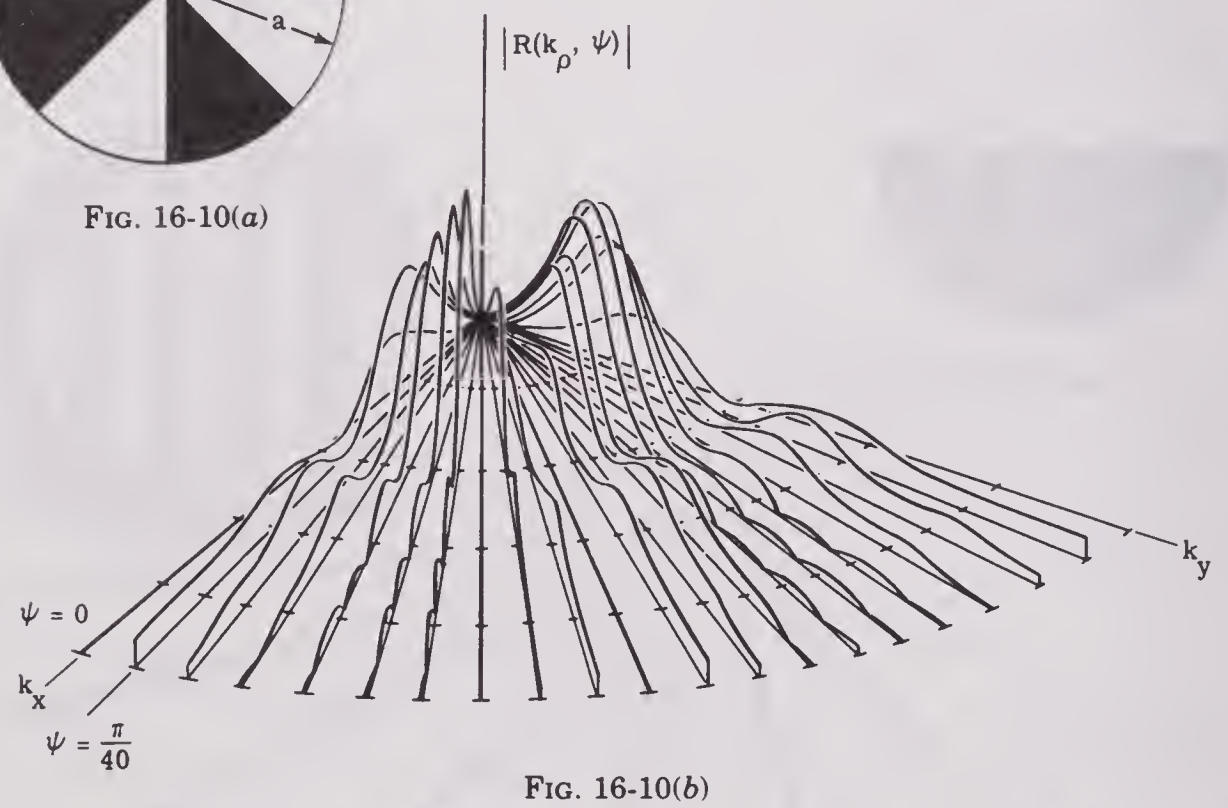


FIG. 16-10(*b*)

Chapter 17

CONTROL SYSTEMS

K. R. Morris

The University of Michigan

CONTENTS

17.1. Linear Systems 662
17.1.1. Basic Definitions 662
17.1.2. Determination of Transfer Functions 664
17.1.3. Methods of Analyzing Linear Systems 668
17.1.4. System Types and Performance 685
17.2. Sampled-Data Systems 695
17.2.1. Basic Definitions 695
17.2.2. Determination of Transfer Functions 700
17.2.3. Methods of Analyzing Sampled-Data Systems 701
17.2.4. Types of Sampled-Data Systems 703
17.3. Nonlinear Systems 703
17.3.1. Basic Definitions 703
17.3.2. Methods of Analyzing Nonlinear Systems 705
17.3.3. Specific Solutions 712
17.4. Design Methods 718
17.4.1. Gain Adjustment 718
17.4.2. Cascade or Series Compensation 718
17.4.3. Root-Locus Method 722
17.4.4. Optimum Transient Response Behavior for Torque-Saturated Systems 724
17.4.5. Miscellaneous Comments on Design and Compensation 724
17.4.6. Statistical Design 725

17. Control Systems

17.1. Linear Systems

17.1.1. Basic Definitions

BASIC SYSTEM. The basic system to be considered is shown in Fig. 17-1. Most systems, however complex, can be reduced to this basic form by applying the rules and techniques given in Sec. 17.1.2.

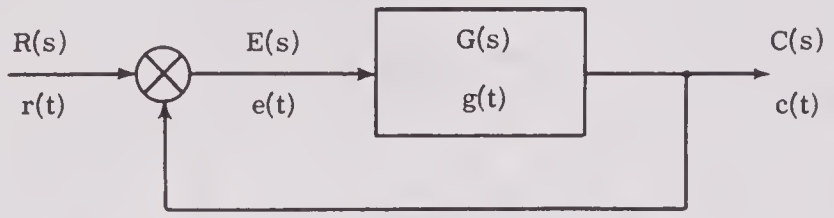


FIG. 17-1. Basic feedback control system.

In general, throughout this chapter, functions represented by lower case letters are functions of the time variable t; capital letters are used to represent the Laplace transform of the lower-case letter function, and are functions of the complex variable $s = \sigma + j\omega$.

TYPICAL TEST SIGNALS

(1) The Delta function, $\delta(t - t_0)$, $u_0\ (t - t_0)$, is zero for all time except t_0, when its height is infinite and its area is unity; it is approximated by a short rectangular pulse of unit area.

(2) The unit step, $h(t - t_0)$, $u_1\ (t - t_0)$, is zero for all $t < t_0$ and one for all $t > t_0$; it is usually specified to be one-half at $t = t_0$. It is approximated by fast-rising voltage (relay or switch with mercury wetted contacts, semiconductor switch, etc.).

(3) The unit ramp, $\omega(t - t_0)t$, $u_2\ (t - t_0)$, is zero for all $t < t_0$, increasing with unity slope for $t > t_0$.

(4) The unit parabola, $\alpha(t - t_0)t^2$, $u_3(t - t_0)$, is zero for all $t < t_0$, increasing at t^2 for $t > t_0$.

TIME DESCRIPTIONS OF SYSTEM BEHAVIOR

(1) The time constant is the time it takes the system to reach $(1/e)A$, when the input is $Ah(t)$ (see Fig. 17-2).

(2) Rise time, t_r, is the time it takes the system to move from $0.1A$ to $0.9A$, when the input is $Ah(t)$ (see Fig. 17-2).

(3) Time delay, t_d, is the time interval between the time of application of input, $Ah(t)$, to the time the system reaches $0.5A$ (see Fig. 17-2).

(4) Overshoot is $(A_p - A)/A$, with input $Ah(t)$ and a peak value of output A_p. Overshoot is measured as a fractional or absolute value, or as a percent (see Fig. 17-2).

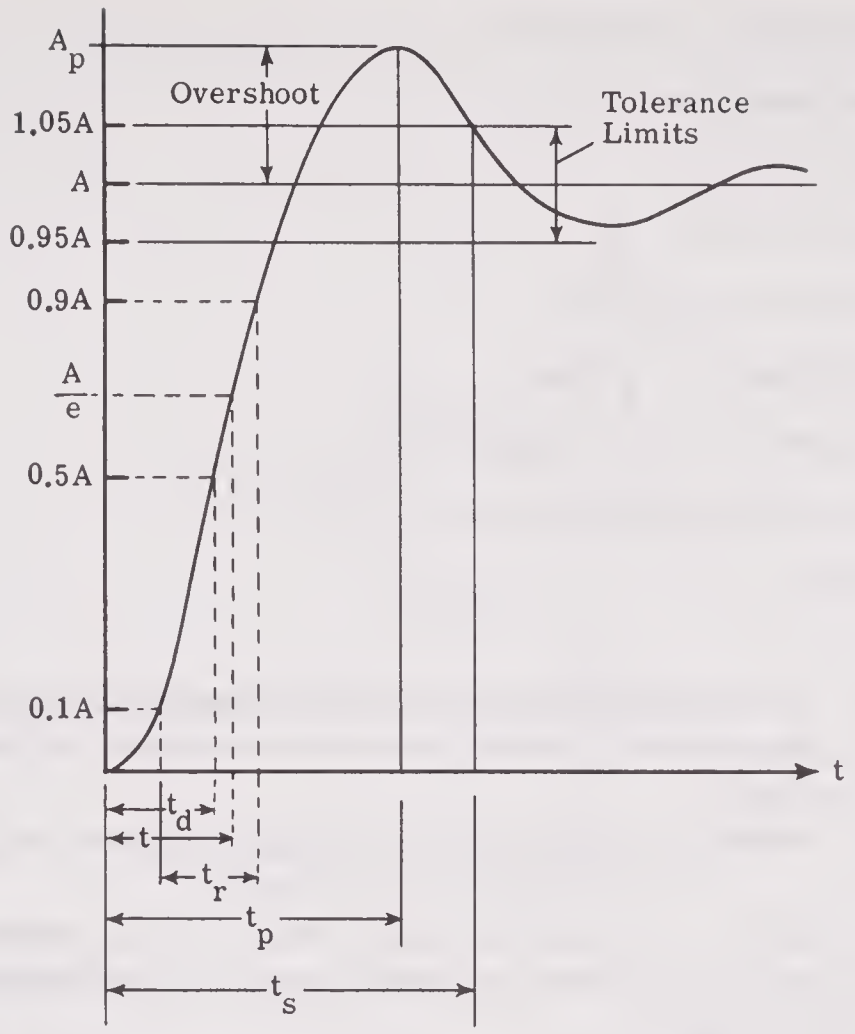


FIG. 17-2. Time descriptions of system behavior.

(5) Time of response, or setting time, t_s, is the time required for the system output (with input $Ah(t)$) to reach and stay within a certain tolerance (e.g., 5%) of the final value (see Fig. 17-2).

(6) Time to peak, t_p, is the interval between application of input $Ah(t)$ and peak output.

(7) The position or displacement error is given by $e = r - c$ (r and c defined as in Fig. 17-1).

(8) The velocity or rate error is given by $\dot{e} = \dot{r} - \dot{c}$. $\left(\dot{x} = \dfrac{dx}{dt}, \text{ where } x \text{ is any variable.}\right)$

(9) The acceleration or parabolic error is given by $\ddot{e} = \ddot{r} - \ddot{c}$. $\left(\ddot{x} = \dfrac{d^2x}{dt^2}\right)$.

WEIGHTING FUNCTION. The weighting function $g(t)$ is the response of a system to a unit impulse input.

TRANSFER FUNCTION. Transfer functions may be defined as follows.

(1) The Laplace transform of weighting function

$$G(s) = L\ \{g(t)\}$$

where $L\{g(t)\}$ = Laplace transform of $g(t) = G(s)$.

(2) The ratio of Laplace transforms of input and output

$$G(s) = \frac{L\{c(t)\}}{L\{r(t)\}} = \frac{C(s)}{R(s)}$$

(3) The frequency response function when it exists (i.e., for stable systems) [1].

The basic forms of transfer functions are:

(*a*) Gain or sensitivity $G(s) = K$

(*b*) Differentiator $G(s) = \tau s$

(*c*) Integrator $G(s) = (\tau s)^{-1}$

(*d*) First-order lead $G(s) = 1 + \tau s$

(*e*) First-order lag $G(s) = (1 + \tau s)^{-1}$

(*f*) Second-order lead $G(s) = s^2 + 2\zeta\omega_n s + \omega_n{}^2$

(*g*) Second-order lag $G(s) = (s^2 + 2\zeta\omega_n s + \omega_n{}^2)^{-1}$

(*h*) Time delay $G(s) = e^{-t_d s}$

STABILITY

Absolute Stability. If the output of a linear system is bounded for any bounded input, the system is said to be absolutely stable (we exclude neutrally stable systems since in practice they do not exist).

Relative Stability. Relative stability is the change necessary to make a system unstable.

MINIMUM PHASE SYSTEM. A minimum phase system is one whose characteristic equation has roots with only negative real parts, hence the minimum phase shift for a given amplitude characteristic.

GAIN MARGIN. Gain margin is the additional gain necessary to make the system unstable.

PHASE MARGIN. Phase margin is the additional phase shift necessary to make the system unstable.

CROSSOVER FREQUENCY. The crossover frequency ω_c, or the frequency to which the loop is closed, is the frequency at which system gain is 0 db.

17.1.2. Determination of Transfer Functions. Transfer functions can be determined from signal flow graphs, block diagrams, Bode plots, and Laplace transforms of transient signals.

17.1.2.1. *Signal Flow Diagrams* [2]

INTRODUCTION. Signal flow diagrams are directly applicable to circuits (electrical, mechanical, etc.) and furnish a method for obtaining the transfer function. They can also be applied to block diagrams for simplification.

The circuit equations are

$$\sum_{i=0}^{n} a_{ij}x_i = 0 \qquad j = 1, 2, 3,\ldots$$

The equations can be rewritten

$$x_j = \sum_{i=0}^{n} t_{ij}x_i \qquad j = 1, 2, 3,\ldots$$

$$t_{ij} = \begin{cases} a_{ij} & \text{when} \quad j \neq i \\ a_{ij} - 1 & \text{when} \quad j = i \end{cases}$$

Figure 17-3 shows the generalized diagram if each x_i is a node, and if n is 3. The transfer function of the entire circuit can be found by applying signal flow rules and diagram simplification.

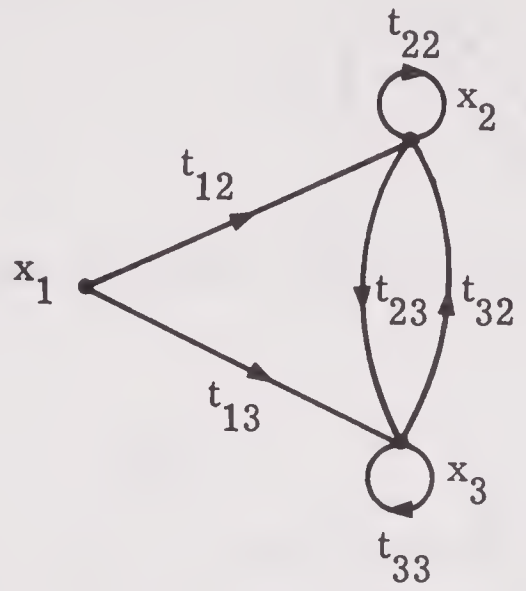


FIG. 17-3. General signal flow diagram for a 3-terminal network.

SIGNAL FLOW RULES

(*a*) Signals travel only in the direction of arrows.

(*b*) A signal is multiplied by the transmittance of the branch t_{ij}.

(*c*) The value of the signal at any node is the sum of the signals entering the node.

(*d*) The signal value of any node is sent along all branches leaving the node.

RULES FOR SIMPLIFYING SIGNAL FLOW DIAGRAMS

(*a*) Parallel paths can be combined by summing transmittances (Fig. 17-4).

(*b*) Series paths can be combined by multiplying transmittances (Fig. 17-5).

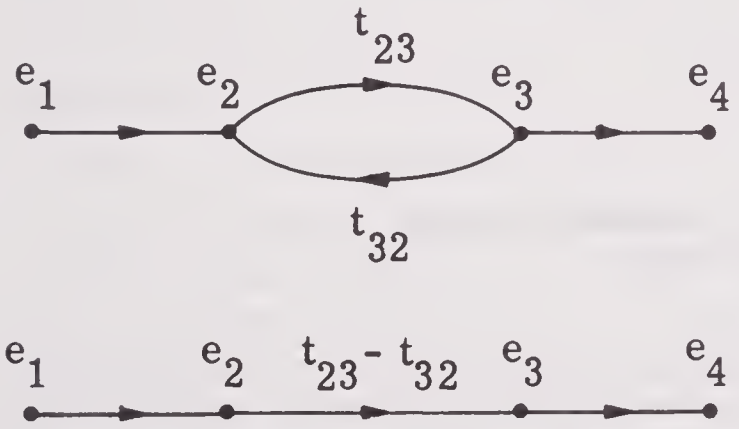


FIG. 17-4. Example of parallel path combination.

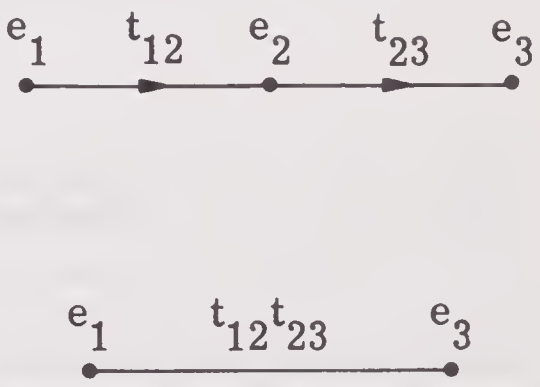


FIG. 17-5. Example of series path combination.

(*c*) Terminal and source ends of branches can be moved. To move the terminal end of a branch t_{31} (see Fig. 17-6), a new node (e_1') must be defined, at which all incoming branches (except the one to be moved) terminate. e_1' is coupled to the old node (e_1) by transmittance 1. A branch from the original node (e_1) to any node (e_5) must

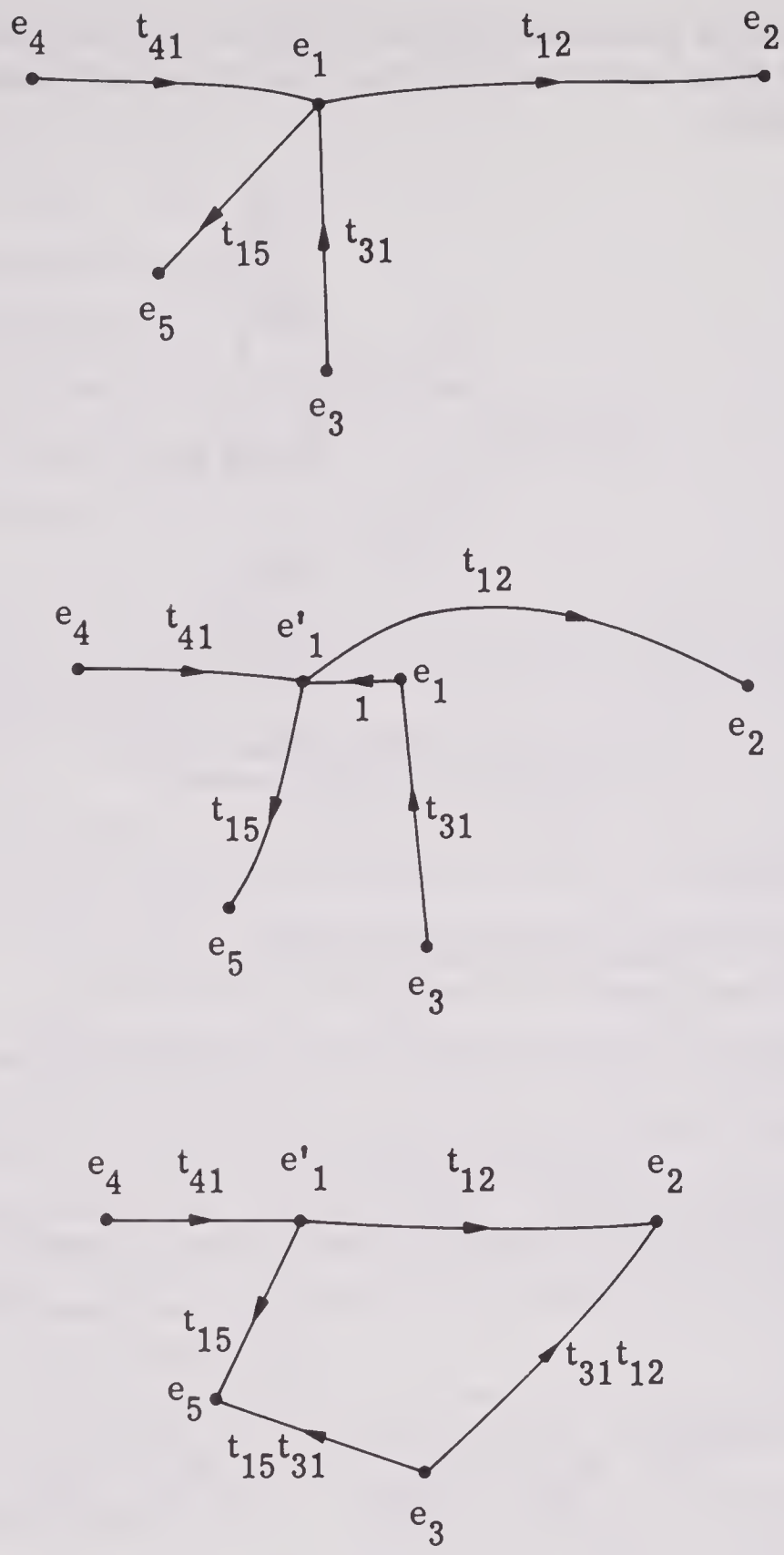


FIG. 17-6. Example of procedure for moving the terminal end of a branch.

have its source moved to e_1' and must receive a branch from the source node (e_3) of the moved branch; this new branch will have transmittance $t_{15}t_{31}$. The terminal end of the branch is moved from e_1 to e_2, multiplying its transmittance by that of the path along which it moved (i.e., $t_{31}t_{12}$).

To remove the source end of a branch from e_3 (Fig. 17-7), each source which had a branch terminating at e_3 must now have a branch terminating at e_2. For example, for e_4 (see Fig. 17-7) the transmittance is $t_{43}t_{32}$, etc.

(*d*) Self loops can be eliminated by multiplying the incoming branch transmittances by $1/(1 - t_{kk})$ (Fig. 17-8).

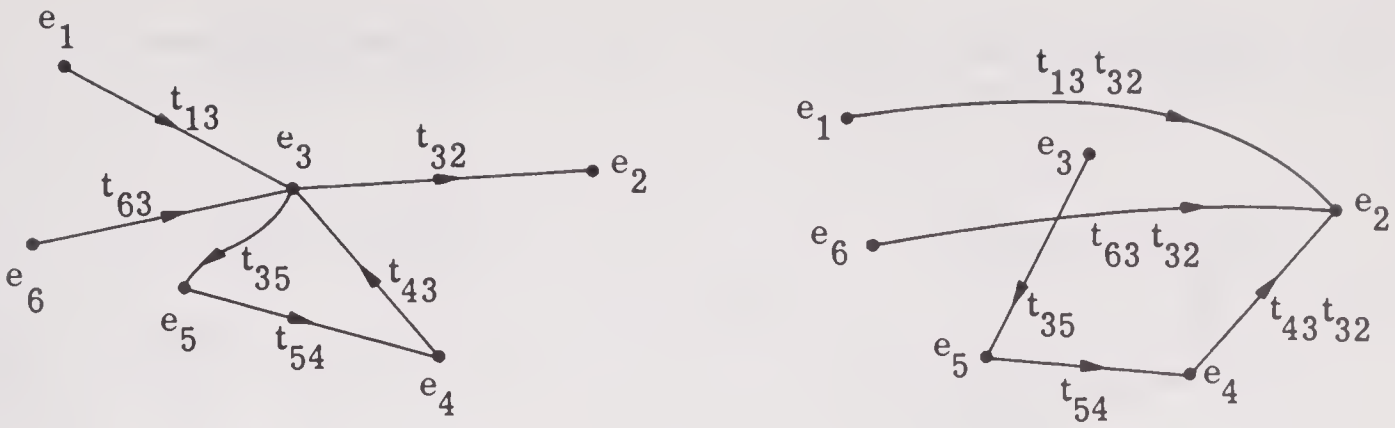


FIG. 17-7. Example of procedure for removing source end of a branch.

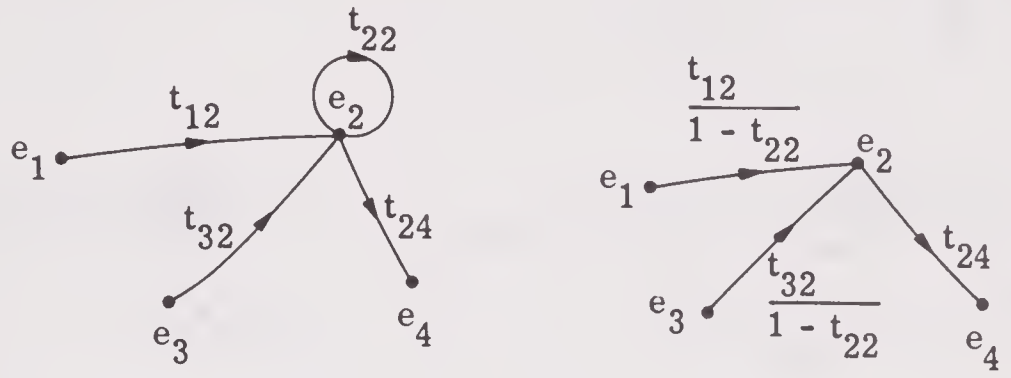


FIG. 17-8. Example of the elinination of a self loop.

17.1.2.2. *Block Diagrams* [3]. Basic rules for the simplification of block diagrams are illustrated in Fig. 17-9, where the contents of each box is the transfer function of that box.

17.1.2.3. *Bode Diagrams.* Experimental curves are obtained which relate system attenuation and phase to frequency (i.e., frequency response curves). Bode's method (see 17.1.3.5) is used to draw approximate asymptotes to the attenuation curve; the phase plot can be used to determine "break" points. The approximate curve can be corrected by adjusting ζ in quadratic terms (see 17.1.3.1 and 17.1.3.5). Further corrections, such as Linville's method [2], are available.

17.1.2.4. *Transient Response Analysis*

GENERAL METHOD. Apply a transient $e(t)$ and record the open-loop system output $c(t)$ (some useful inputs are found in Sec. 17.1.1). Then the ratio of the Laplace transforms of response and input, $G(s) = C(s)/E(s)$, yields the transfer function $G(s)$. For complicated $c(t)$, one can use approximations made from a series of simple functions (such as steps, triangles, impulses, rectangles, etc.), and then find the transform of each; the total transform is the sum of the transforms of the individual components.

GUILLEMIN'S IMPULSE METHOD [2]. An impulse is applied to the system to obtain $g(t)$; $g(t)$ is then approximated by a sequence of straight lines, impulses, parabolic curves, cubic curves, or other higher powers of t, yielding $g^*(t)$. $g^*(t)$ is differentiated until only impulse functions remain; then

$$G^*(j\omega) = \sum_{k=1}^{n} \frac{a_k e^{-j\omega t_k}}{(j\omega)^{\nu}}$$

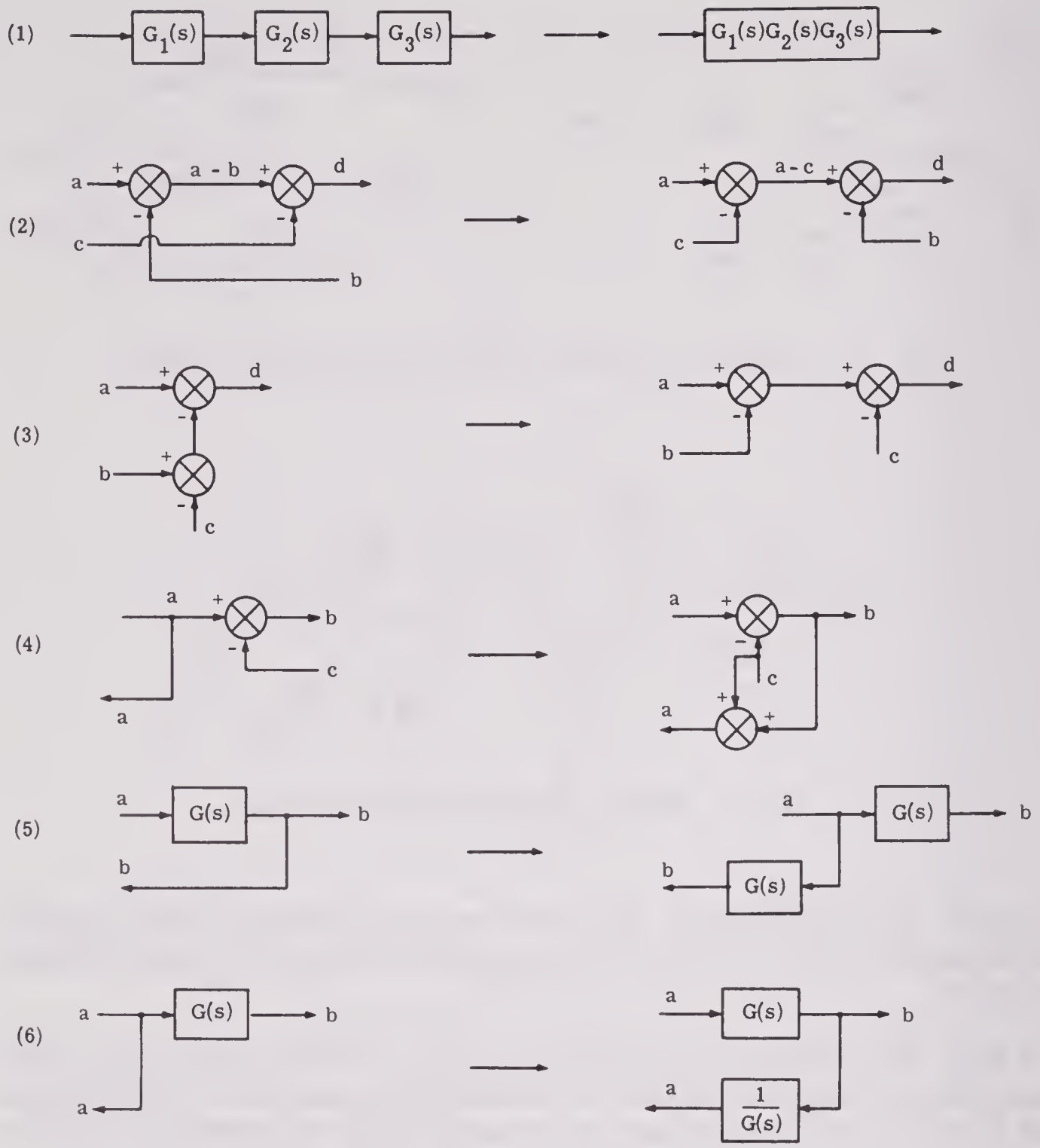


FIG. 17-9. Basic rules for block diagram simplification.

where $G^*(j\omega)$ is the approximate transform of $g(t)$

a_k is the amplitude of the kth impulse

t_k is the time of the kth impulse

ν is the number of derivatives taken to obtain impulses

n is the total number of impulses used.

Approximating $\dot{g}(t)$ by straight lines is equivalent to approximating $g(t)$ by parabolic segments.

17.1.3. Methods of Analyzing Linear Systems. A linear system can be analyzed by considering the differential equations describing it, by its weighting function, by Nyquist methods, by the Bode method, by log modulus plots, and by root locus methods.

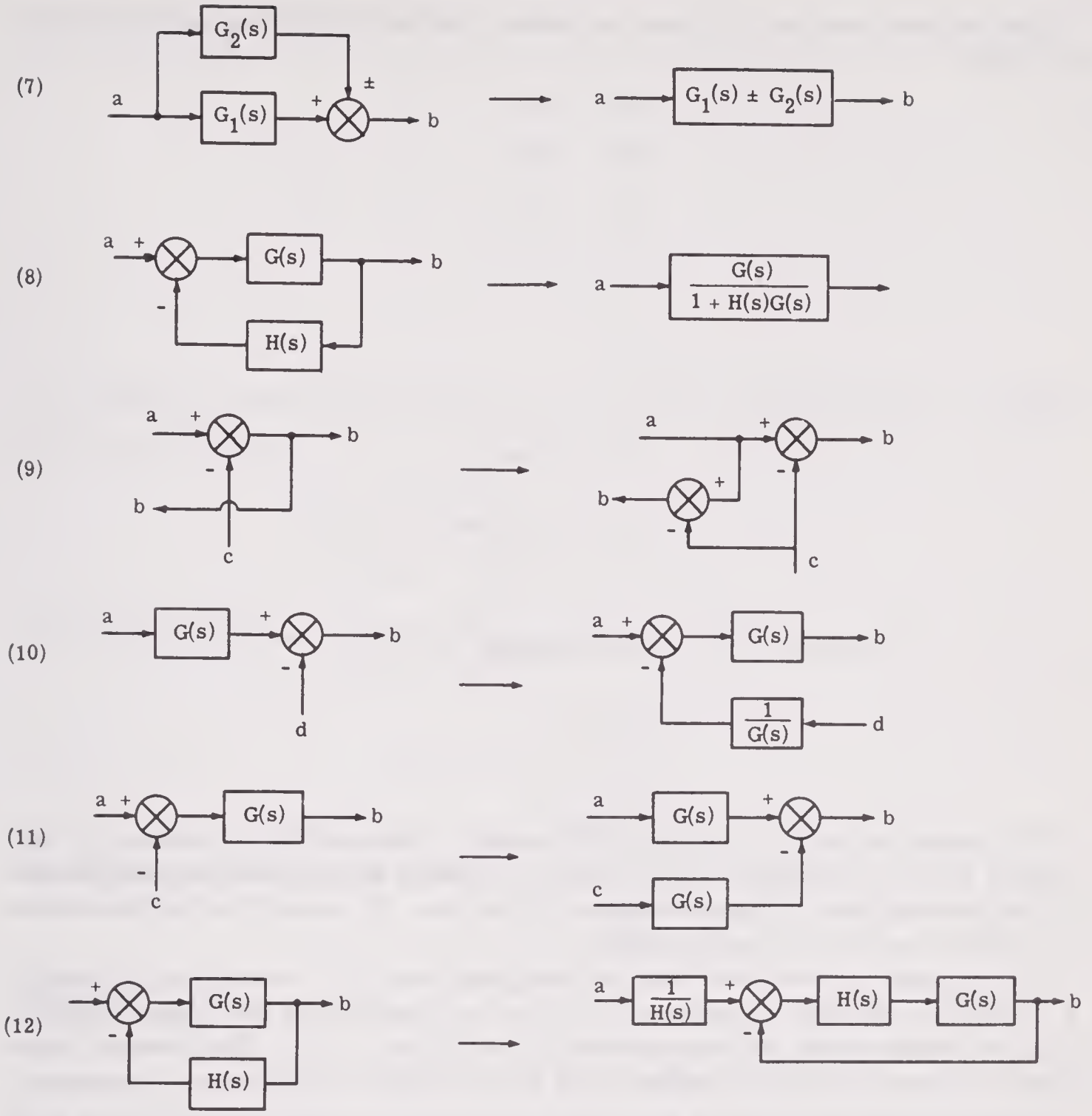


FIG. 17-9 (*Continued*). Basic rules for block diagram simplification.

17.1.3.1. *Differential Equation Method*

REPRESENTATION OF THE SYSTEM. Sum the torques (forces, voltages, etc.) on the output and write the equation of motion (where $p = d/dt$),

$$C(p) = R(p)$$

The characteristic equation is $C(p) = 0$.

ABSOLUTE STABILITY ANALYSIS. If the characteristic equation has any derivatives missing (lower than the order of the equation), or if the coefficients of the characteristic equation are not all of the same sign, the system is unstable. If the roots of the characteristic equation have any positive real parts, the system is unstable. A useful test for positive real roots is Routh's Stability Test I [4]:

Write the characteristic equation

$$C(s) = \sum_{i=0}^{n} a_i s^{n-i} = 0$$

Form an array from the coefficients, as follows, continuing until the last row has only one entry:

$$\begin{array}{llll} a_0 & a_2 & . & . \\ a_1 & a_3 & . & . \\ b_1 & b_3 & . & . \\ c_1 & c_3 & . & . \\ . & & & \\ . & & & \\ . & & & \end{array}$$

where

$$b_1 = -\det \begin{vmatrix} a_0 & a_2 \\ a_1 & a_3 \end{vmatrix}$$

$$b_3 = -\det \begin{vmatrix} a_0 & a_4 \\ a_1 & a_5 \end{vmatrix}$$

$$c_1 = -\det \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}$$

$$\begin{array}{l} . \\ . \\ . \end{array}$$

The number of zeros in the right-half s-plane is indicated by the number of sign changes in the first column. Work can be simplified by dividing each element of a row by the absolute value of the first element of that row. If an entry in the first column is 0, replace the zero by ϵ and continue.

If all elements of a row are 0, there are imaginary roots [5]. An auxiliary polynomial is formed from the last non-zero row; the order is n (the order of the original equation is m, the number of the last non-zero row is r, and $n = m - r + 1$). The powers of s contained in this polynomial will either be all even or all odd. Differentiate and proceed.

If the characteristic equation is written in terms of K (a gain constant), limits on K for stability can be determined such that the sign requirement is satisfied.

Example

$$(s-2)\,(s+3)\,(s+5)\,(s_2+9)$$

$$s^5 + 6s^4 + 8s^3 + 24s^2 - 9s - 270$$

Row			
1:	1	8	−9
2:	1	4	−45
	[6]	[24]	[−270]
3:	1	9	
	[4]	[36]	
4:	−1	−9	
	[−5]	[−45]	
5:	−2	0	$\therefore$ form $-s^2 - 9 \rightarrow -2s$
	[0]		

$\therefore$ one root with positive real part, one pair of complex, conjugate roots.

An alternate form is Routh's Test II:

Form the array:

$$\begin{array}{ccccccc} a_1 & a_0 & 0 & 0 & 0 & \cdots \\ a_3 & a_2 & a_1 & 0 & 0 & \\ a_5 & a_4 & a_3 & a_2 & a_1 & \\ \cdot & & & & & \\ \cdot & & & & & \\ \cdot & & & & & \end{array}$$

For stability, all determinants formed by the elements in the upper left corner must have the same sign, i.e.,

$$\begin{vmatrix} a_1 & a_0 \\ a_3 & a_2 \end{vmatrix} \quad \begin{vmatrix} a_1 & a_0 & 0 \\ a_3 & a_2 & a_1 \\ a_5 & a_4 & a_3 \end{vmatrix}, \ldots$$

STEADY-STATE BEHAVIOR $c_s(t)$. The system equation can be written as

$$c(t) = r(t) \left[\frac{\sum_{i=0}^{n} a^i p^i}{-p^N \sum_{i=0}^{m} b^m p^m} \right]$$

Form the power series

$$c(t) = r(t) \left[p^{-N} \left(1 - \frac{1}{b_0 - a_0} p - \ldots \right) \right]$$

Example

$$c(t) = \frac{1}{\tau p + 1} r(t) = (1 - \tau p + \tau^2 p^2 - \ldots) r(t)$$

then

$$\text{if } r(t) = \delta(t),\ c_s(t) = 0$$

$$\text{if } r(t) = Ah(t),\ c_s(t) = A$$

$$\text{if } r(t) = \omega t,\ c_s(t) = \omega t - \tau\omega$$

Alternately, for $r(t) = Ah(t)$, set all derivatives in the characteristic equation to zero.

Example

$$(p^2 + 4p + 253)c(t) = r(t)(p^3 + 32)$$

$$c_s(t) = \frac{32A}{253}$$

TRANSIENT BEHAVIOR $c_T(t)$. Assume a solution of the characteristic equation

$$c_T(t) = \sum_{i=1}^{n} C_i e^{\lambda_i t}$$

where the λ_i are the eigenvalues of the characteristic equation and the C_i are determined from the initial conditions.

For many systems the performance is dominated by two complex poles (no zeros), such that

$$G(s) \approx (s^2 + 2\zeta\omega_n s + \omega_n^2)^{-1}$$

Then the system has the following characteristics:

(*a*) ζ is the damping factor

(*b*) ω_n is the undamped natural frequency

(*c*) $\omega_n(1 - \zeta^2)^{1/2} = \omega_d$, the damped frequency

(*d*) time to peak $t_p = \pi/\omega_d$

(*e*) settling time (to $\pm 2\%$) $t_s \approx 4/\zeta\omega_n$

(*f*) overshoot $= \exp(-\zeta\omega_n t_p)$

(*g*) rise time (10 to 90%) $t_r \approx 2.2/\zeta\omega_n$

(*h*) logarithmic decrement δ

$$\delta = 1/n \ln x_0/x_n T = 2\pi n\zeta/(1 - \zeta^2)^{1/2}$$

where x_0 is the amplitude of the first overshoot (measured from the final value) and x_{nT} is the amplitude of the signal after n cycles of oscillation. ζ is the damping ratio of a second-order system. This is a valid measure only for lightly damped systems ($\zeta < 0.1$).

17.1.3.2. *Weighting Function Method*

Representation of the System. The weighting function $g(t)$ is the output obtained when a unit impulse is applied to the input. Step inputs, ramp inputs, etc., can be used if the output is differentiated (once for step inputs, twice for ramp inputs, etc.).

Absolute Stability Analysis. The James-Weiss stability criterion [1] states that for absolute stability

$$\int_0^\infty |g(\tau)|\, d\tau < \infty$$

Time Behavior from Weighting Function. Determine the output of the system either graphically or analytically by using

$$c(t) = \int_0^\infty e(t)\, g(t - \tau)\, d\tau$$

This method is not as useful for determining damping factors, frequency response, etc.; but it does give $c(t)$ directly.

17.1.3.3. *Nyquist Method* [6]

Representation of the System. Write the open-loop transfer function as

$$KG(s) = KA(s)/B(s)$$

On Fig. 17-10 locate all poles and zeros (singular points) on the s-plane which lie on the $j\omega$ axis. Select a clockwise contour enclosing all poles and zeros in the right-half plane (a semicircle at infinity is chosen). This contour should bypass each pole and zero on the $j\omega$ axis by a small counter-clockwise semicircle.

Evaluate $KG(j\omega) = |KG(j\omega)| \angle \phi$ for enough values of ω that $KG(j\omega)$ can be plotted as a function of ϕ (the Nyquist or polar plot in the KG plane). The plot is symmetrical around the real axis. As ω travels around the semicircle enclosing the singular points on the $j\omega$ axis, the Nyquist plot contour moves through an angle of approximately $\pm 180n°$ at very large radius. (n = order of singularity; + for poles, − for zeros.) Often, $G(j\omega)$ is plotted instead of $KG(j\omega)$ (G-plane Nyquist plot) (see Fig. 17-11).

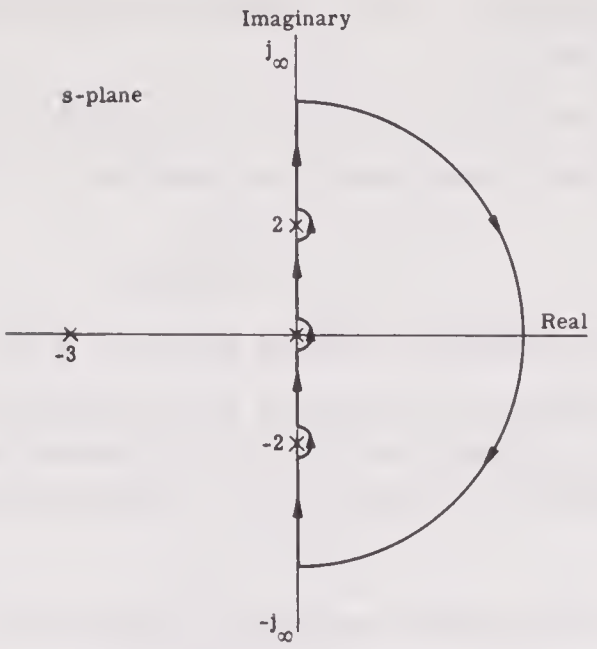


FIG. 17-10. s-plane contour for

$$G(s) = \frac{1}{s(s+3)(s^2+4)}.$$

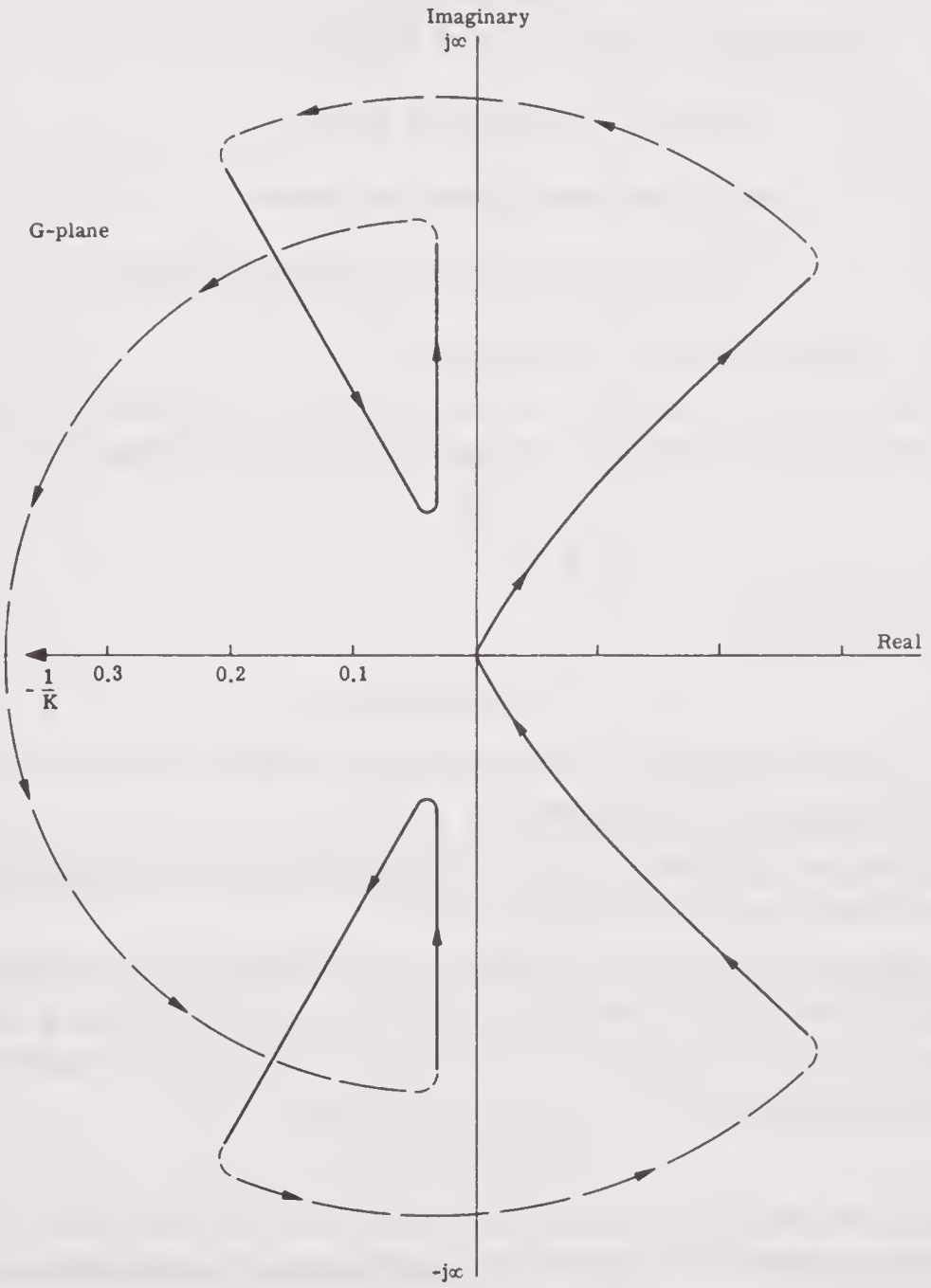


FIG. 17-11. G-plane contour (Nyquist plot) for

$$G(s) = \frac{1}{s(s+3)(s^2+4)}$$

ABSOLUTE STABILITY ANALYSIS. First, determine the number of clockwise encirclements (N) of the critical point.

(a) In the KG plane, the open-loop critical point is the origin.

(b) In the KG plane, the closed-loop critical point is -1, since

$$G_c(s) = \frac{KG(s)}{1 + KG(s)}$$

(c) In the G plane, the closed-loop critical point is $-1/K$

Encirclement can be determined by tracing around the KG (or G) contour in a clockwise direction; if the critical point is on the right, it is inside the contour, hence encircled.

Next, determine the number of poles in the right-half plane (zeros of $B(s)$). Then

$$Z = P - N$$

where P = the number of right-half-plane poles circled by the s-plane

N = the number of encirclements of critical points

Z = the number of right-half-plane zeros in the system

STEADY-STATE BEHAVIOR. The steady-state behavior is determined by the manner in which the G or KG contour behaves as $\omega \to 0$.

(a) If $\lim_{\omega \to 0} G$ = closed plot, the system had displacement error

(b) If $\lim_{\omega \to 0} G = -j\infty$, the system has velocity error

(c) If $\lim_{\omega \to 0} G = -\infty$, the system has acceleration error

(d) If $\lim_{\omega \to 0} G = +j\infty$, the system has rate-of-acceleration error

(e) Lower-order errors are 0 (see also Sec. 17.1.4).

TRANSIENT BEHAVIOR. M circles (see Fig. 17-12) are the locus of constant ratio of open-loop to closed-loop magnitude. The coordinates of the center are

$$x = \frac{-1}{K} \frac{M^2}{M^2 - 1} \qquad y = 0$$

The radius is given by

$$r = \frac{1}{K} \frac{M}{M^2 - 1}$$

(a) M_p is the peak amplitude of the closed-loop response, and $G_c(s)$ is given by the M circle tangent to the contour (Fig. 17-13).

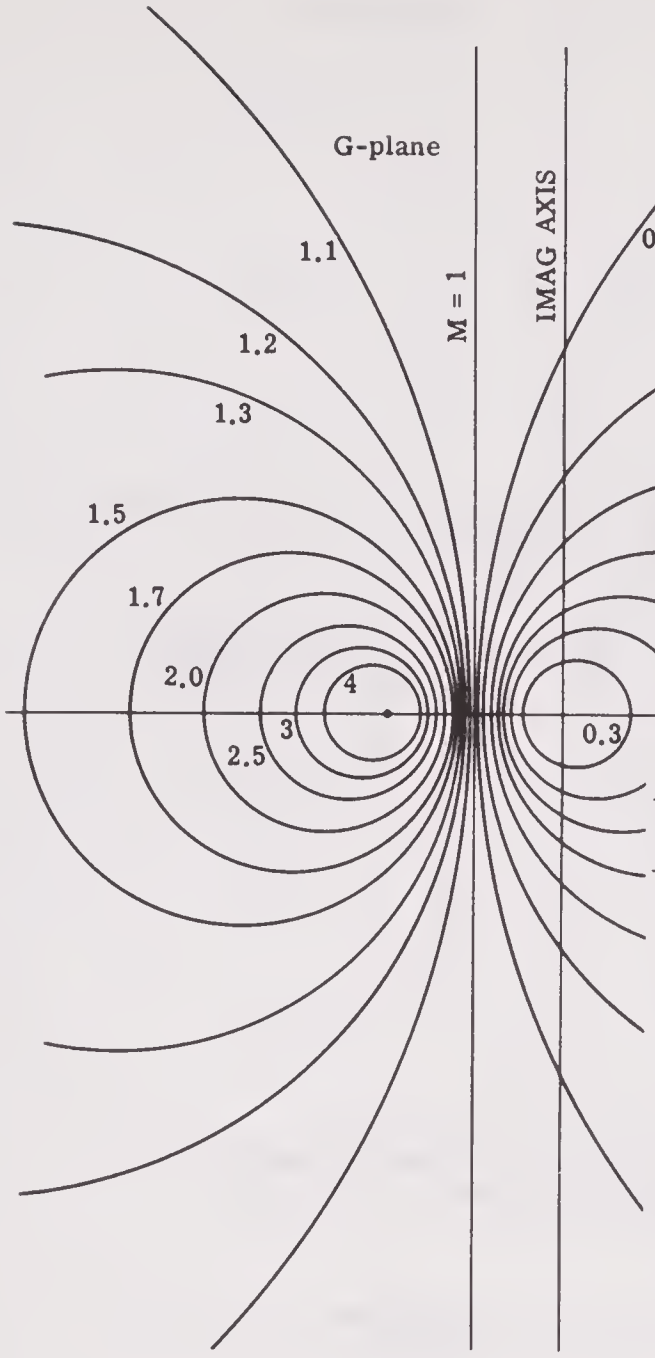
(b) ω_p, the frequency of resonance, is the frequency on the contour at the point of tangency with the M_p circle (Fig. 17-13).

(c) The magnitude of gain at any point of the contour, ω_i, is determined by first drawing an M circle through the point of interest. Then a ψ line is drawn through the origin tangent to the M circle. If ψ is the angle made by the ψ line and the negative real axis, then (see Fig. 17-13)

$$1/M = \sin \psi$$

(d) The gain constant K corresponding to any point of the contour is determined by a line perpendicular to the real axis and passing through the point of tangency of the M circle and the ψ line (see Fig. 17-13).

(e) Frequencies for which the locus is to the right of the line $M = 1$ (circle of infinite radius) are attenuated.



M LOCI FOR POLAR PLOTS [6]

M	Center (x-coordinate)	Radius	$\psi = \sin^{-1} 1/M$, deg
0.5	+0.333	0.67	
0.6	+0.562	0.94	
0.7	+0.960	1.37	
0.8	+1.777	2.22	
0.9	+4.26	4.74	
1.0	∞	∞	90
1.1	−5.77	5.24	65.2
1.2	−3.27	2.73	56.5
1.3	−2.45	1.88	50.2
1.4	−2.04	1.46	45.6
1.5	−1.80	1.20	41.8
1.6	−1.64	1.03	38.7
1.7	−1.53	0.9	36
1.8	−1.47	0.84	33.7
1.9	−1.38	0.73	31.7
2.0	−1.33	0.666	30.0
2.2	−1.26	0.573	27.0
2.4	−1 21	0.505	24.6
2.5	−1.19	0.476	23.6
2.6	−1.17	0.451	22 6
2.8	−1.15	0.41	20.9
3.0	−1.12	0.375	19.5
3.5	−1.10	0.34	16.6
4.0	−1.07	0.266	14.5
4.5	−1.05	0.234	12.8
5.0	−1.04	0.208	11.5

FIG. 17-12. M-circle loci for polar (Nyquist) plots.

The best transient response (best because the transfer function exhibits the least peaking) is given by the K value derived from ψ_{max} (the largest ψ which is still tangent to the locus).

Since the distance and angle from the origin to the locus give $G(j\omega)$ while that from the critical point $-1/K$ gives $+1/K + G(j\omega)$ (Fig. 17-14), the total system response $G_c(j\omega)$ is given by

$$G_c(s) = \frac{KG(s)}{1 + KG(s)}$$

hence the ratio of the two distances gives the magnitude of $G_c(s)$, and the difference between the angles gives the phase shift.

The phase margin is given by the angle formed by the vector of length $1/K$ which intersects the G contour, and the negative real axis. N circles (Fig. 17-15) are circles

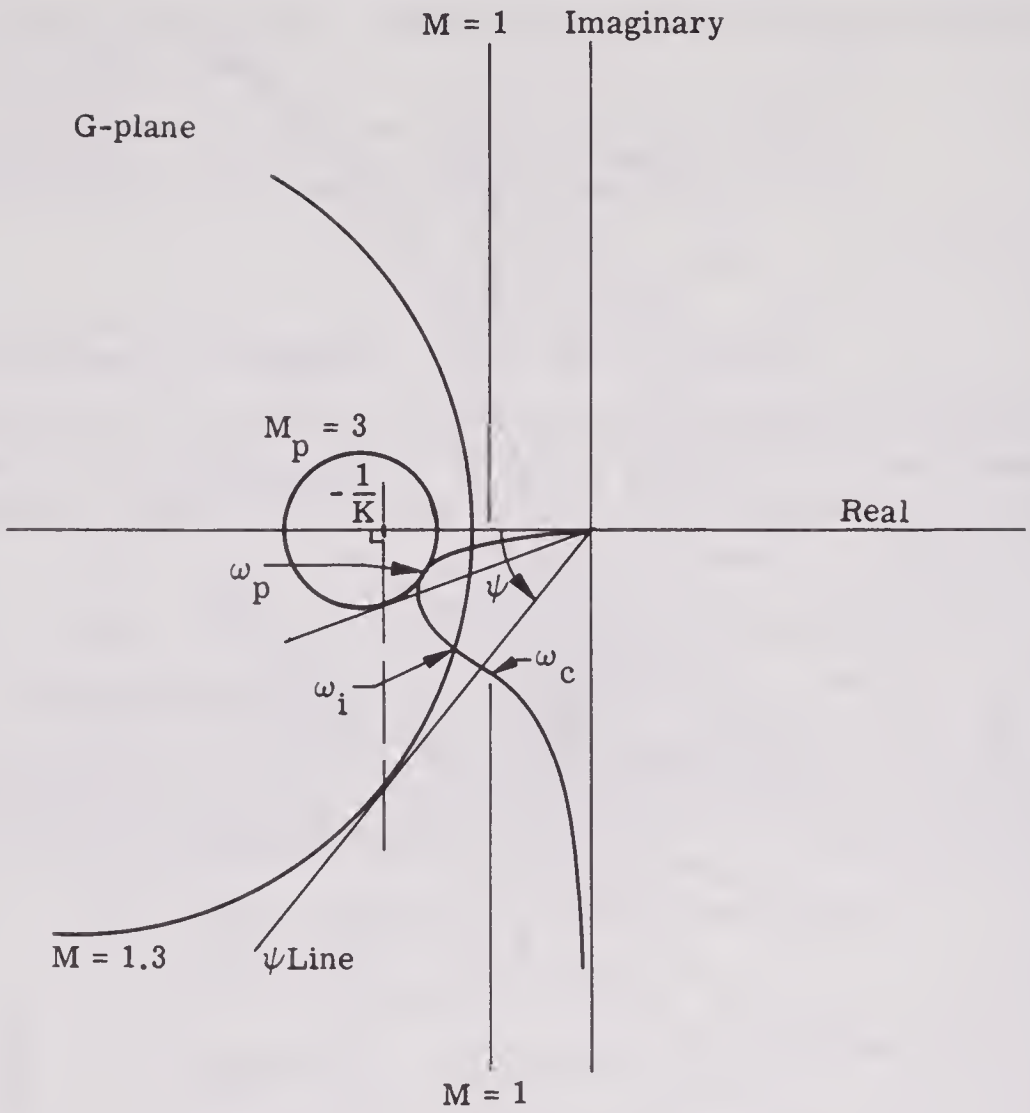


FIG. 17-13. Example of transient behavior determination from *M*-circles.

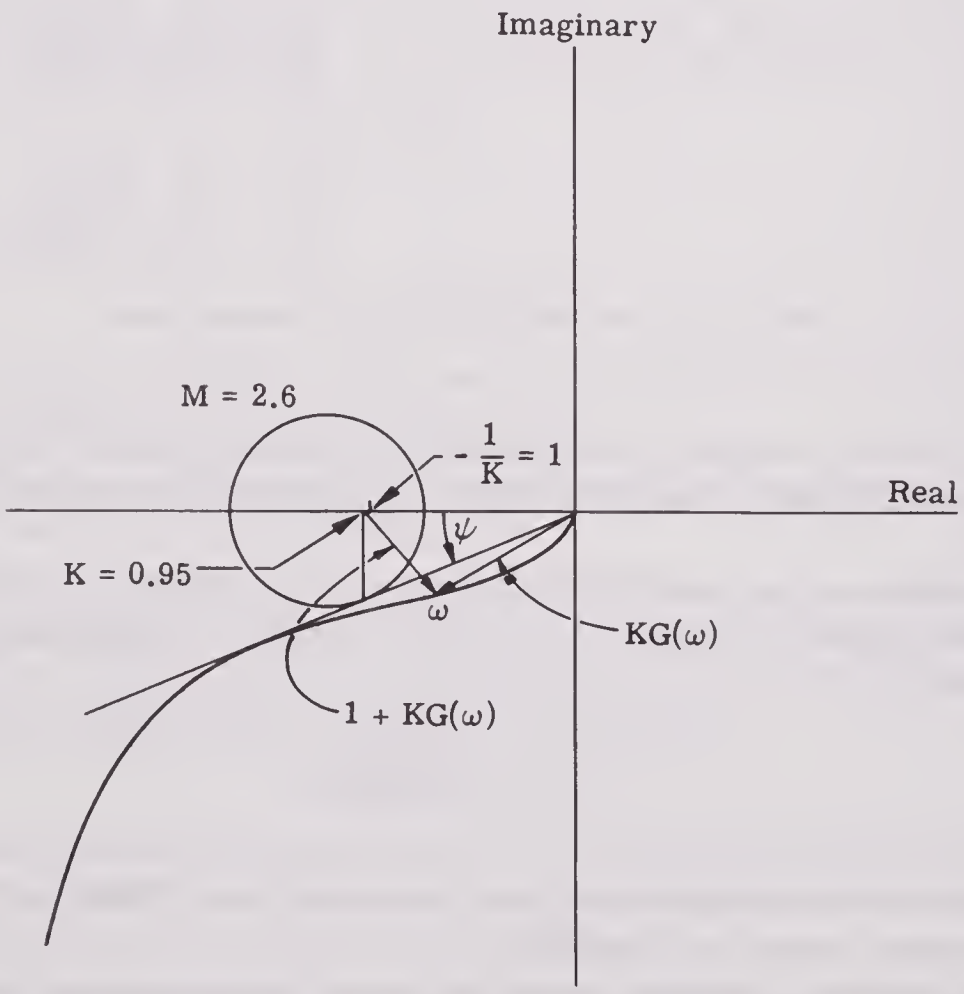
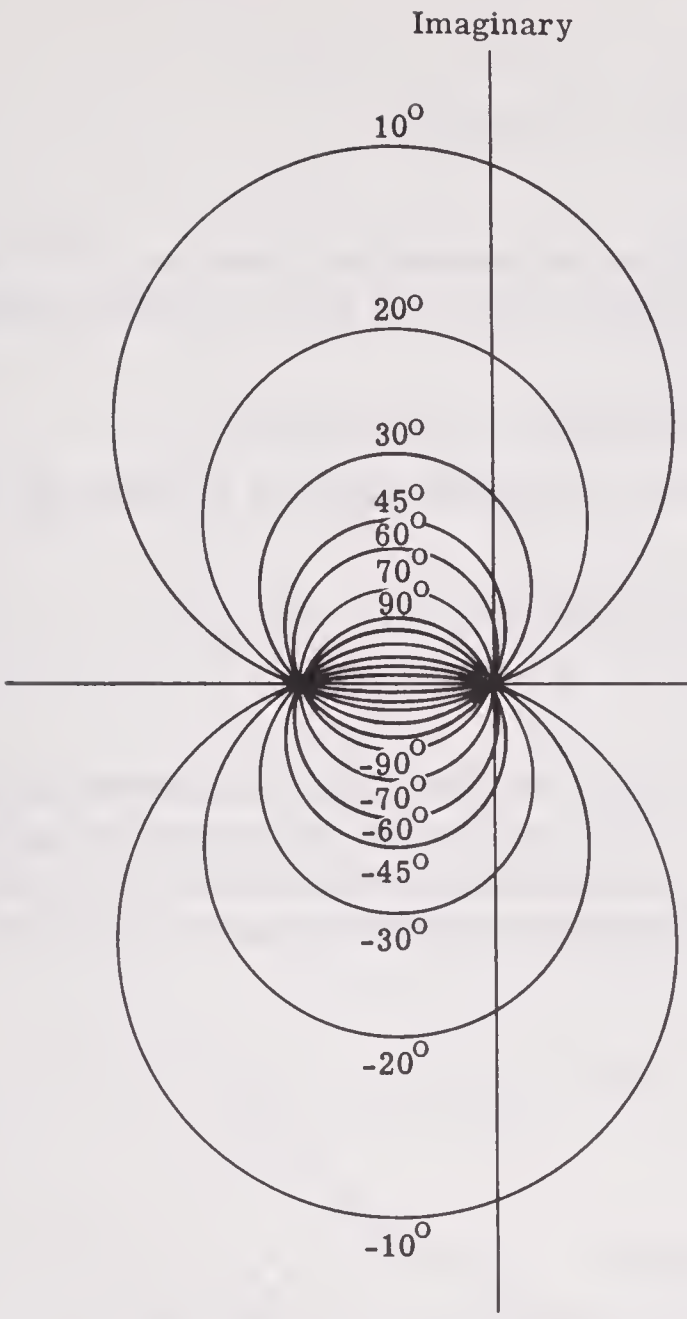


FIG. 17-14. System gain from *M*-circle.



$\theta_o - \theta_c$	Center x = - 1/2 y	Radius
±10°	±2.84	2.88
±20°	±1.37	1.46
±30°	±0.866	1.00
±45°	±0.5	0.707
±60°	±0.29	0.577
±70°	±0.18	0.531
±90°	0	0.500

FIG. 17-15. *N*-circle loci for polar (Nyquist) plots.

of constant phase difference between input and output signals; they are characterized by a center

$$x = -1/2, \qquad y = +1/2N$$

$$N = \tan(\theta_o - \theta_c)$$

where θ_o is the open-loop phase shift

θ_c is the closed-loop phase shift

The radius of the N circles is given by

$$R = \left[\frac{1}{4} + \left(\frac{1}{2N}\right)\right]^{1/2}$$

At the frequency at which a locus crosses an N circle, the difference between input and output phase is given by the value of $\tan^{-1} N$.

17.1.3.4. *Inverse Nyquist Method.* When components $H(s)$ are found in the feedback loop such that

$$G_c(s) = \frac{KG(s)}{1 + KG(s)H(s)}$$

it is convenient to consider

$$[G_c(s)]^{-1} = \frac{1}{KG(s)} + AF(s)H(s)$$

The sum can be performed graphically by vector summation at each ω. M^{-1} circles are concentric about $-1/K$; other characteristics carry over from the Nyquist method in a similar manner.

17.1.3.5. *Bode Method (log Magnitude and Phase vs log Frequency).*

REPRESENTATION OF THE SYSTEM. Write the transfer function in terms of basic forms, i.e.,

$$G(s) = \frac{K(\tau_1 s + 1)\,(\tau_2 s + 1)\,\ldots}{s^N(\tau_a s + 1)\,(\tau_b s + 1)\,(s^2 + 2\zeta\omega_n s + \omega_n{}^2)\,\ldots}$$

Next, construct a straight-line approximation to the Bode diagram by summing these basic components (Fig. 17-16 and 17-17). The diagram can be corrected by applying the curves of Fig. 17-18 and 17-19. An example is shown in Fig. 17-20. If the system is non-minimum phase, the phase portrait will be modified as shown in Fig. 17-18*b*.

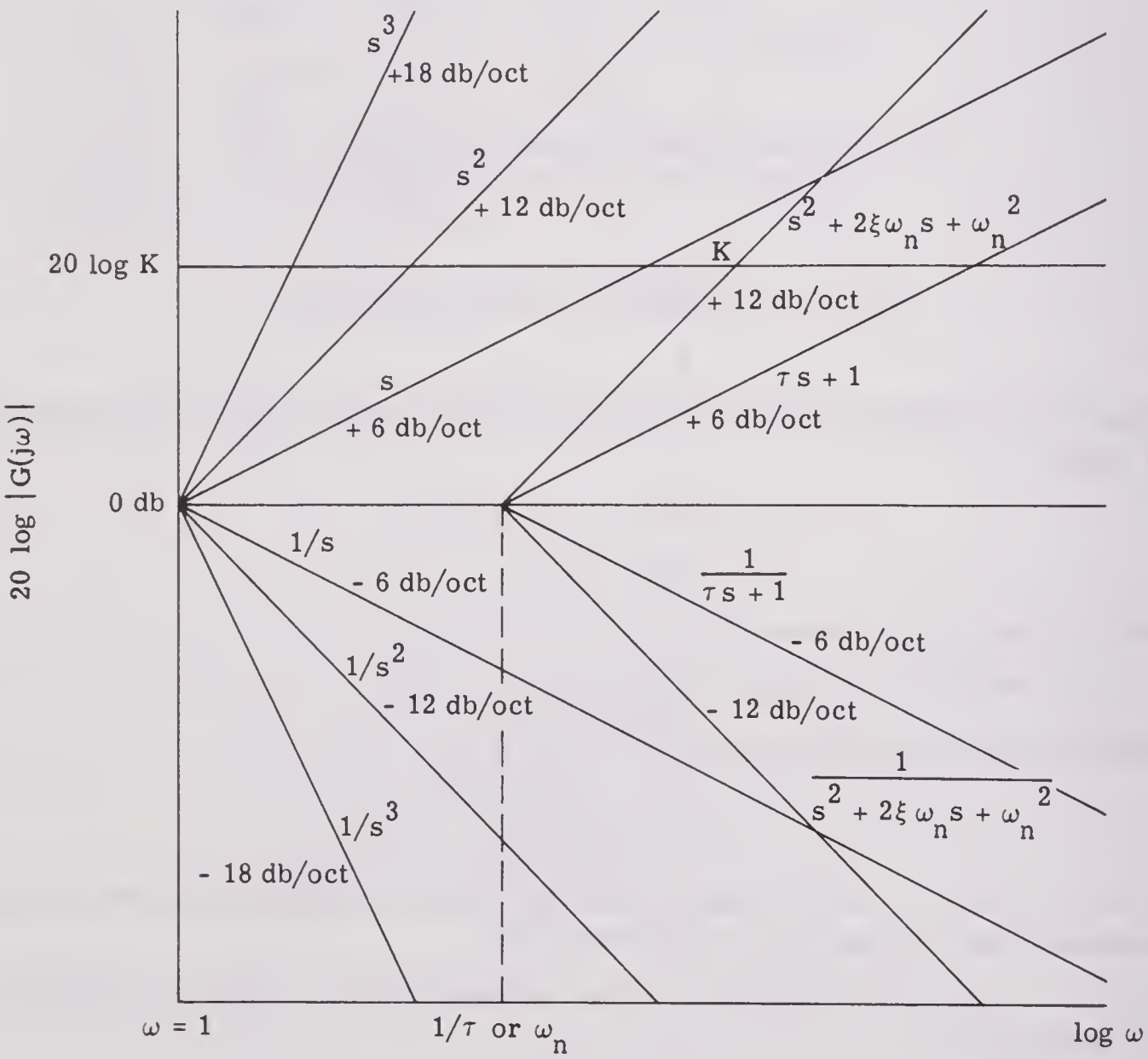


FIG. 17-16. Straight-line approximations to basic forms for Bode plot, $|G(j\omega)|$.

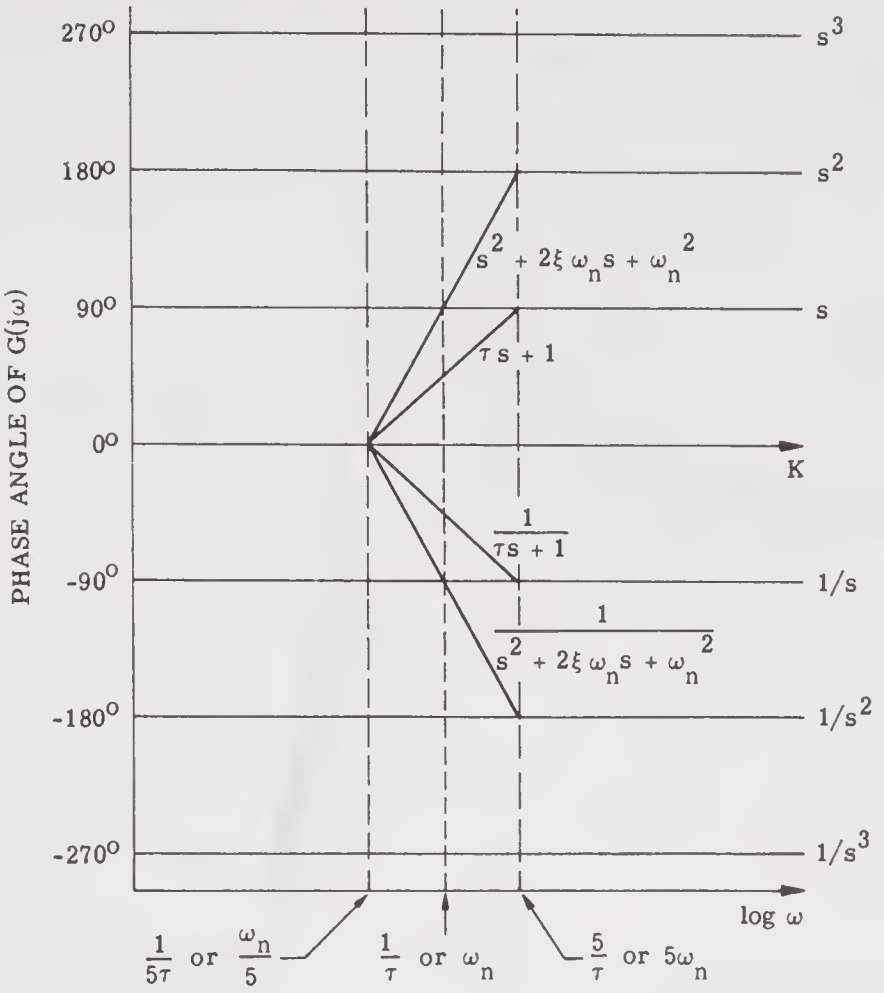


FIG. 17-17. Straight-line approximations to basic forms for Bode plot, $\underline{/G(j\omega)} = \phi$.

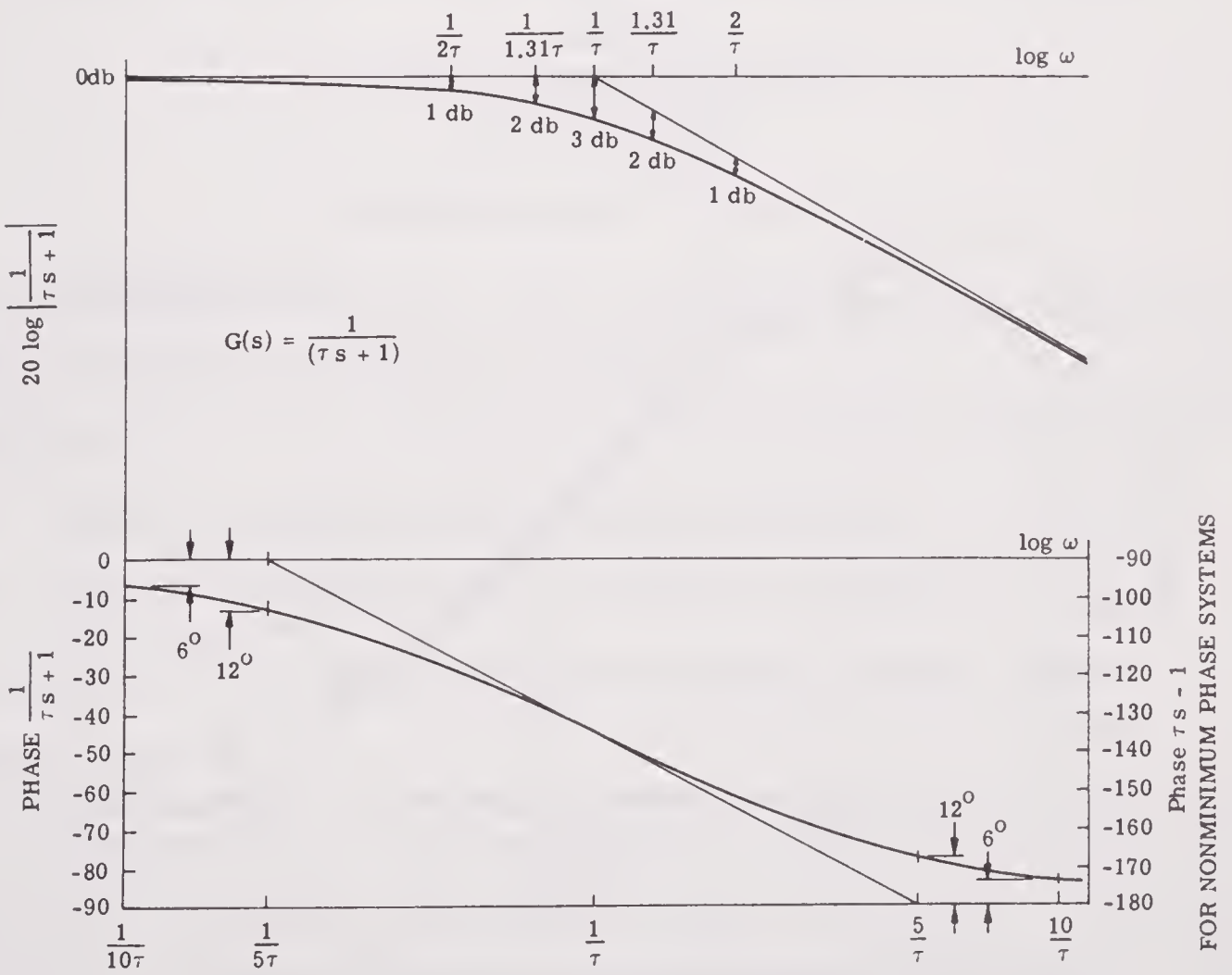


FIG. 17-18. Corrections for Bode plot of first-order basic forms.

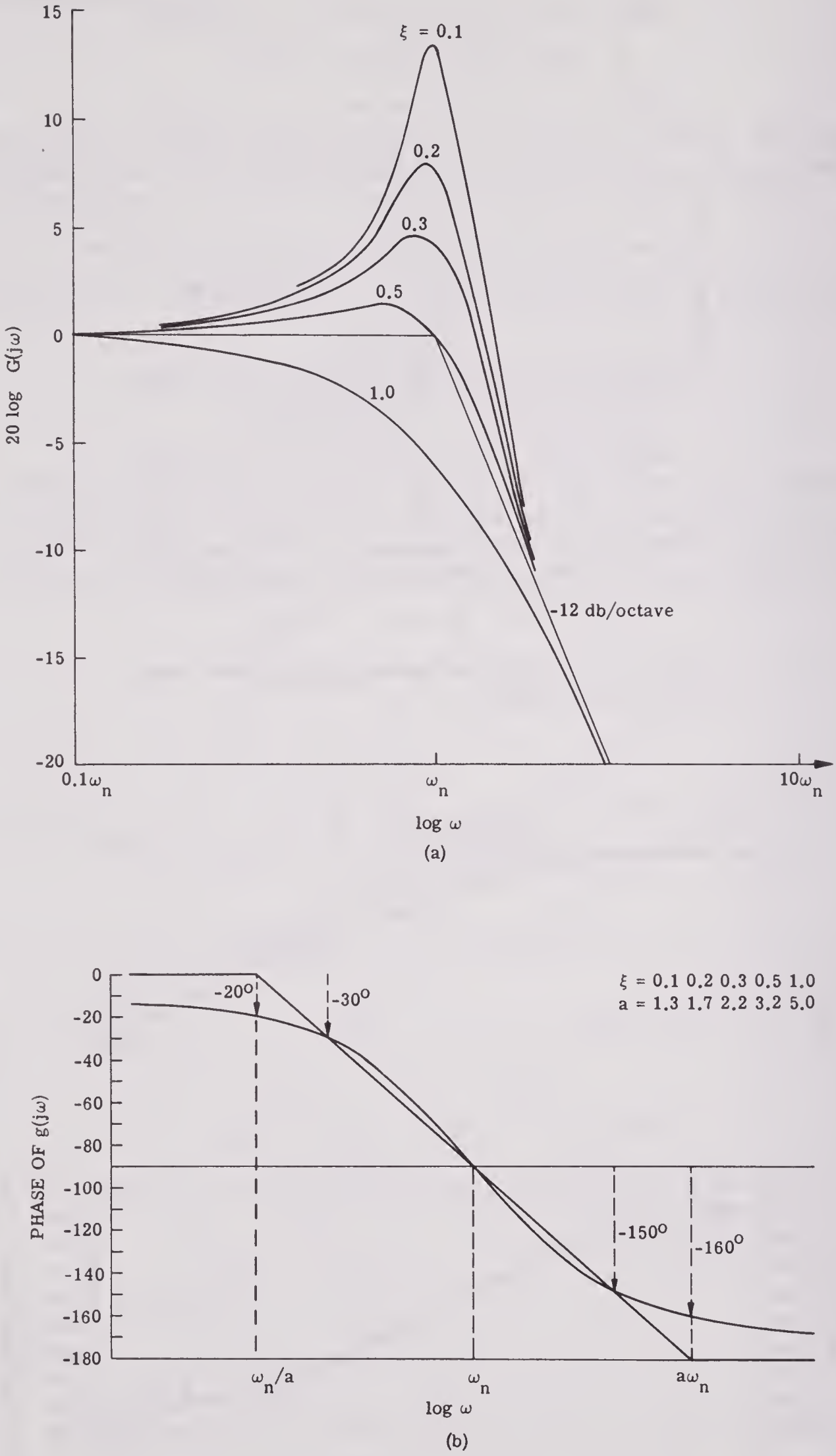
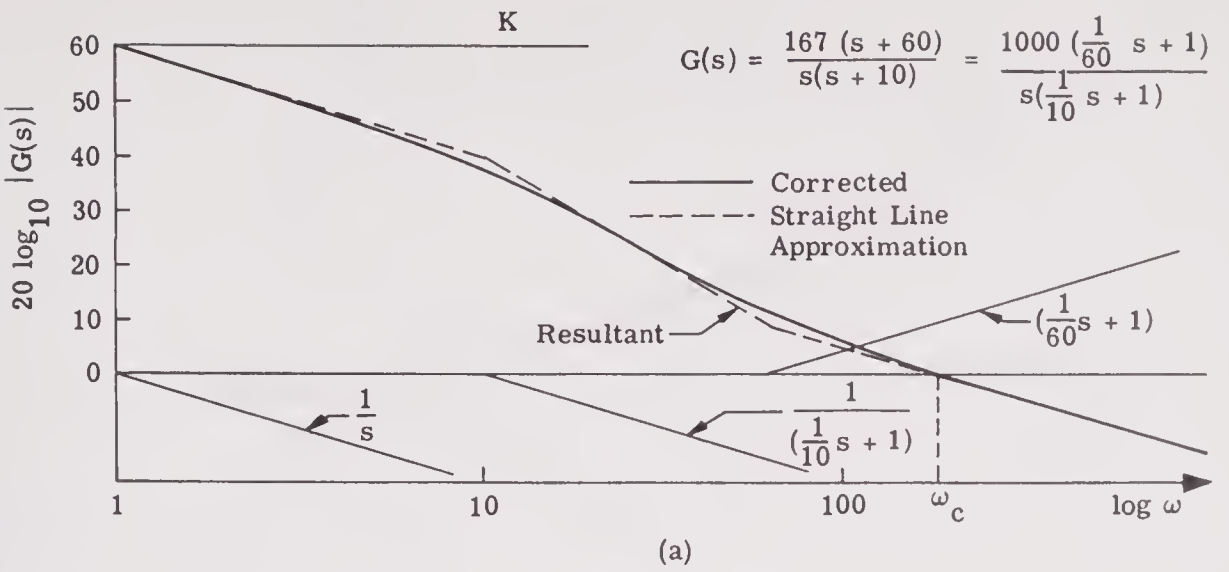
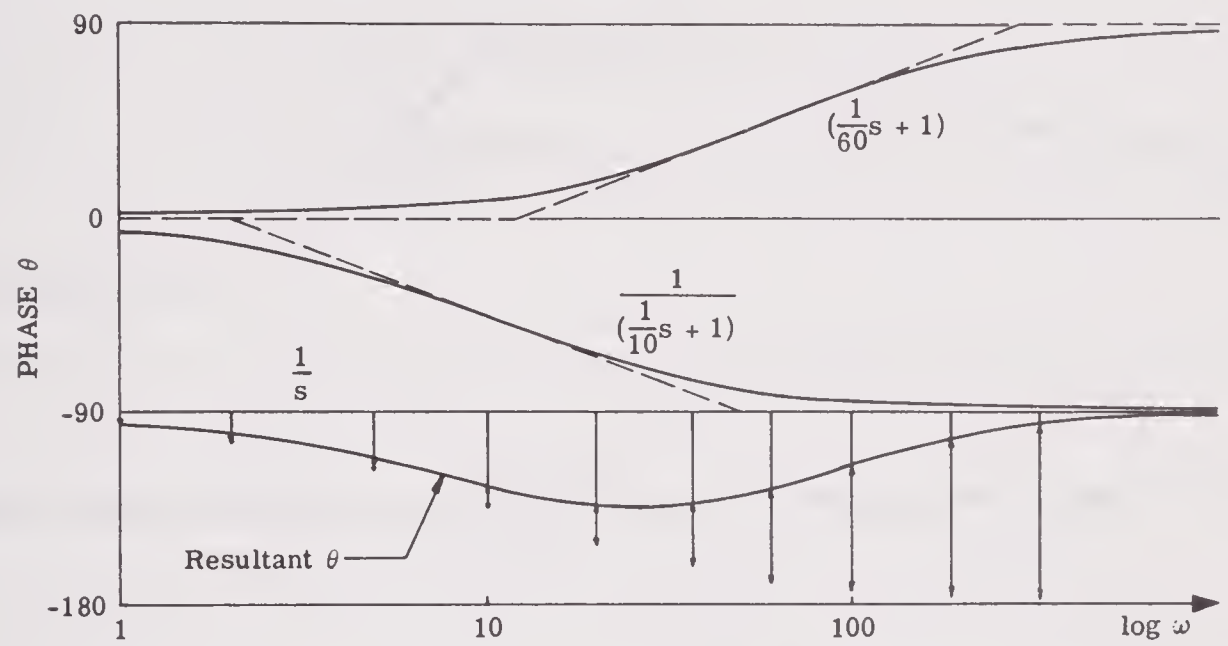


FIG. 17-19. Corrections for Bode plot of second-order basic forms. (*a*) Amplitude correction. (*b*) Phase correction.



(a)



(b)

FIG. 17-20. Example of Bode plot for

$$G(s) = \frac{167(s + 60)}{s(s + 10)}$$

(*a*) Amplitude of $G(j\omega)$. (*b*) Phase of $G(j\omega)$.

ABSOLUTE STABILITY ANALYSIS. If the system gain is above unity (0 db) when the phase angle exceeds 180°, the system is unstable.

STEADY-STATE BEHAVIOR. The steady-state behavior is determined by the slope of the diagram:

(*a*) If $\lim_{\omega \to 0} |G| = 0$ slope, the system has displacement error

(*b*) If $\lim_{\omega \to 0} |G| = -6$ db/octave slope, the system has velocity error

(*c*) If $\lim_{\omega \to 0} |G| = -12$ db/octave slope, the system has acceleration error

See Sec. 17.1.4.

TRANSIENT BEHAVIOR. In general, the greater the phase margin, the smaller the transient (see Fig. 17-21). A totally equalized system is stable for all gain with a slope of −6 db/octave and a phase shift of −90° for $0 < \omega < \omega_c$; the response is almost purely exponential. A partially equalized system is stable for some gains, i.e., the slope is −6 db/octave for ω near ω_c; in this case, the response is somewhat oscillatory. Sometimes a system is referred to by the slopes of the approximate response, *e.g.*, 12-6-12.

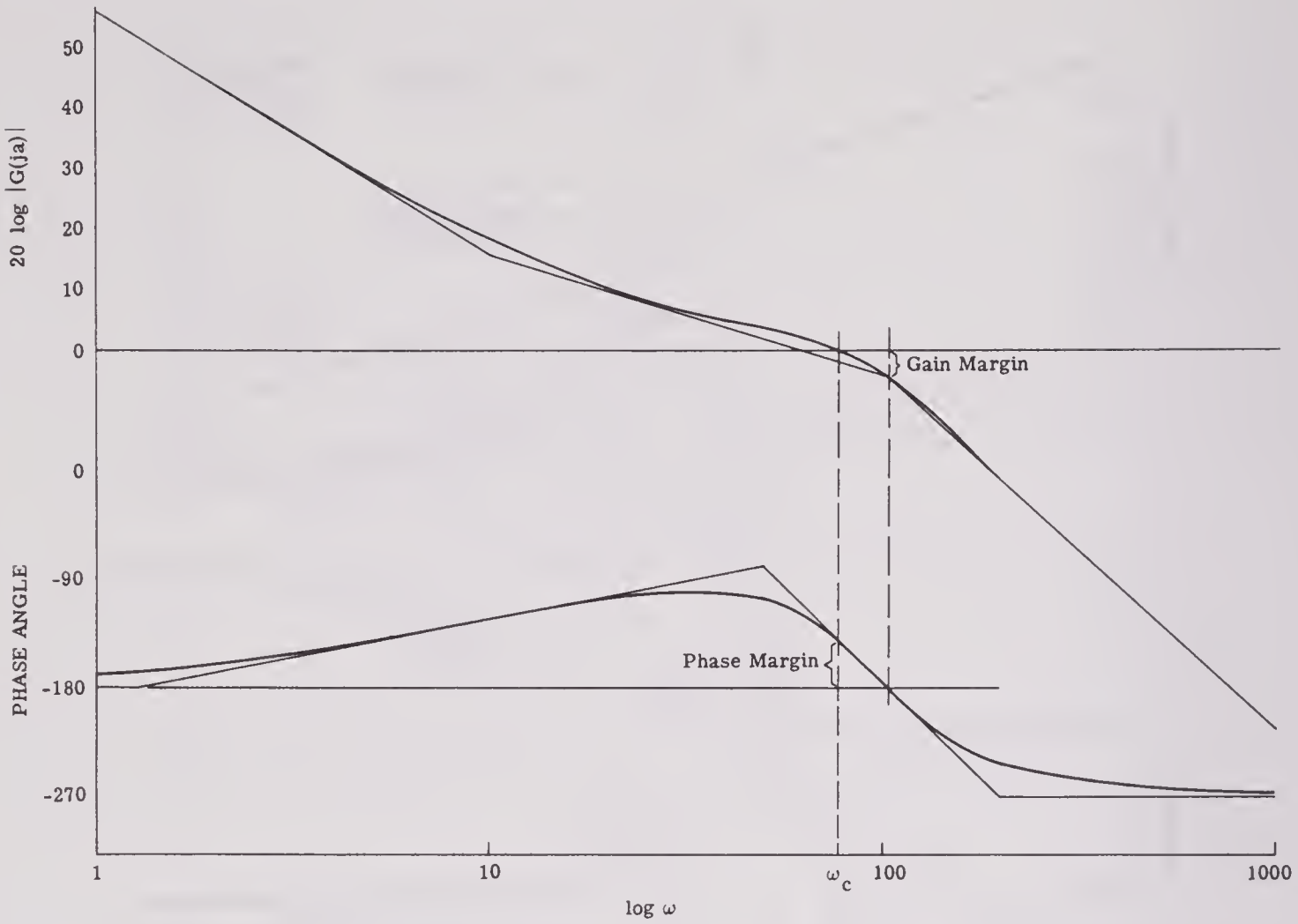


FIG. 17-21. Example of Bode plot for $G(s) = \dfrac{63.1(s+10)}{s^2(s^2+100s+10^4)}$ showing phase and gain margins.

17.1.3.6. *Log Modulus Method (log Magnitude vs Phase)*

REPRESENTATION OF THE SYSTEM. Write the transfer function as

$$KG(s) = \frac{KA(s)}{B(s)} = |G(s)| \angle \phi(s)$$

Plot 20 log $|G(s)|$ *vs* $\phi(s)$; the total system plot is the sum of the angle and magnitude curves for the system components (Fig. 17-22).

ABSOLUTE STABILITY ANALYSIS. If the system response diagram intercepts the $\pm 180°$ axis above 0 db, the system is unstable.

STEADY-STATE BEHAVIOR. Steady-state behavior is determined by the way the contour behaves as $\omega \to 0$. See Sec. 17.1.4.

(*a*) If $\underset{\omega \to 0}{\phi} \to 0°$, the system has displacement error

(*b*) If $\underset{\omega \to 0}{\phi} \to -90°$, the system has velocity error

(*c*) If $\underset{\omega \to 0}{\phi} \to -180°$, the system has acceleration error

TRANSIENT BEHAVIOR. M and N contours (Nichols chart) (Fig. 17-23) can be used as in the Nyquist method; the phase margin is indicated as the contour intercept with the 0-db axis, and the gain margin is indicated as the contour intercept with the 180° axis.

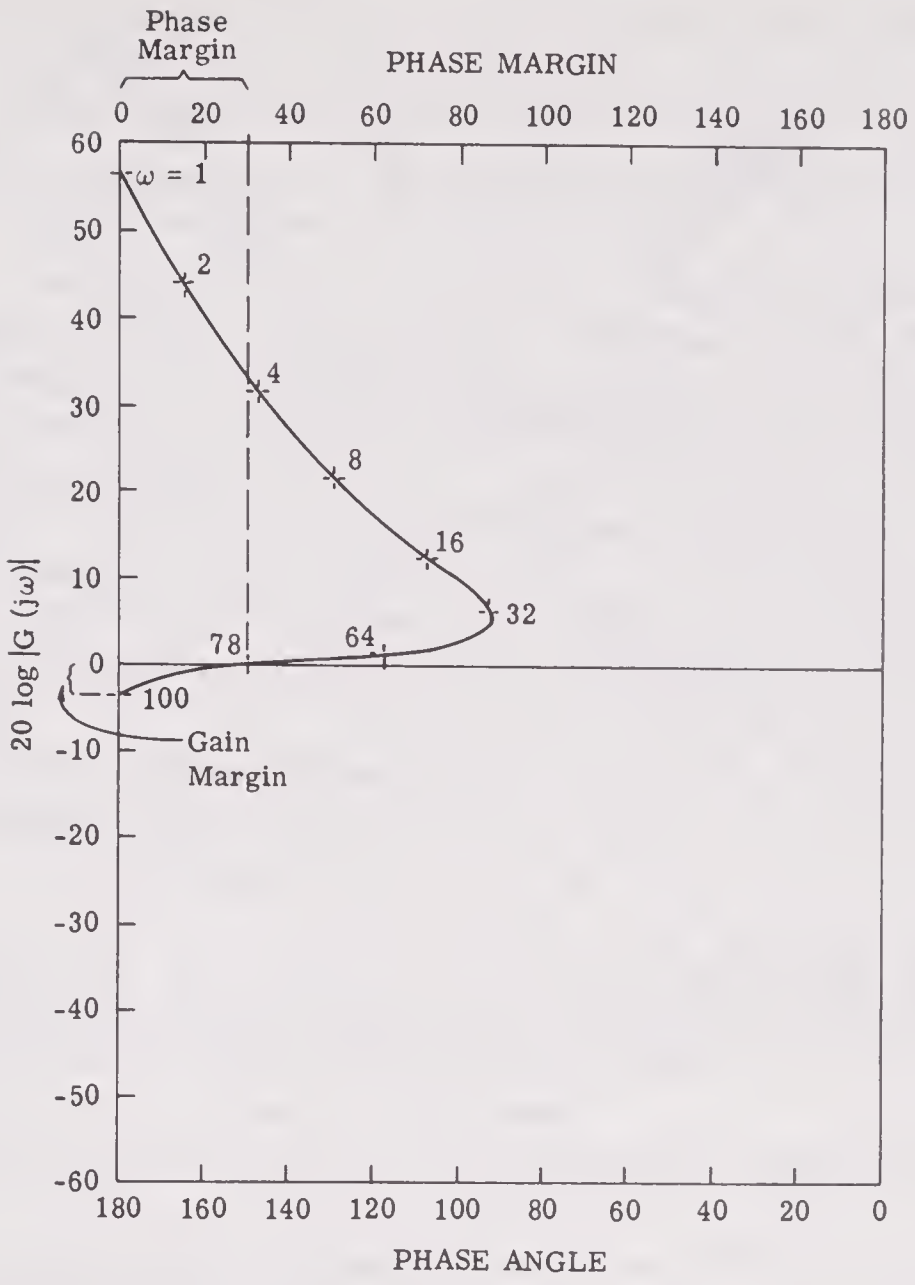


FIG. 17-22. Example of log modulus plot for

$$G(s) = \frac{63.1(s + 10)}{s^2(s^2 + 100s + 10^4)}$$

17.1.3.7. *Root Locus Method*

REPRESENTATION OF THE SYSTEM. A root-locus plot shows the locus of closed-loop poles as the open-loop gain is changed. Consider the transfer function of the open-loop system $G(s) = p(s)/q(s)$. The zeros of $q(s)$ (i.e., poles of $G(s)$) are marked on a diagram as x's, the zeros of $p(s)$ as o's. The rules for construction are as follows:

(a) The loci start at poles.

(b) The loci terminate at zeros.

(c) The number of loci is the number of poles or zeros of $G(s)$.

(d) If the coefficients of p and q are real, complex parts of the loci appear in conjugate pairs.

(e) The loci approach ∞ along asymptote lines at $\pm 180°/n$, $\pm 540°/n$, $\pm 900°/n$, etc., where n is determined by letting s get very large so that only the highest-order terms of p and q remain of importance. Then

$$G(s) \approx K/s^n$$

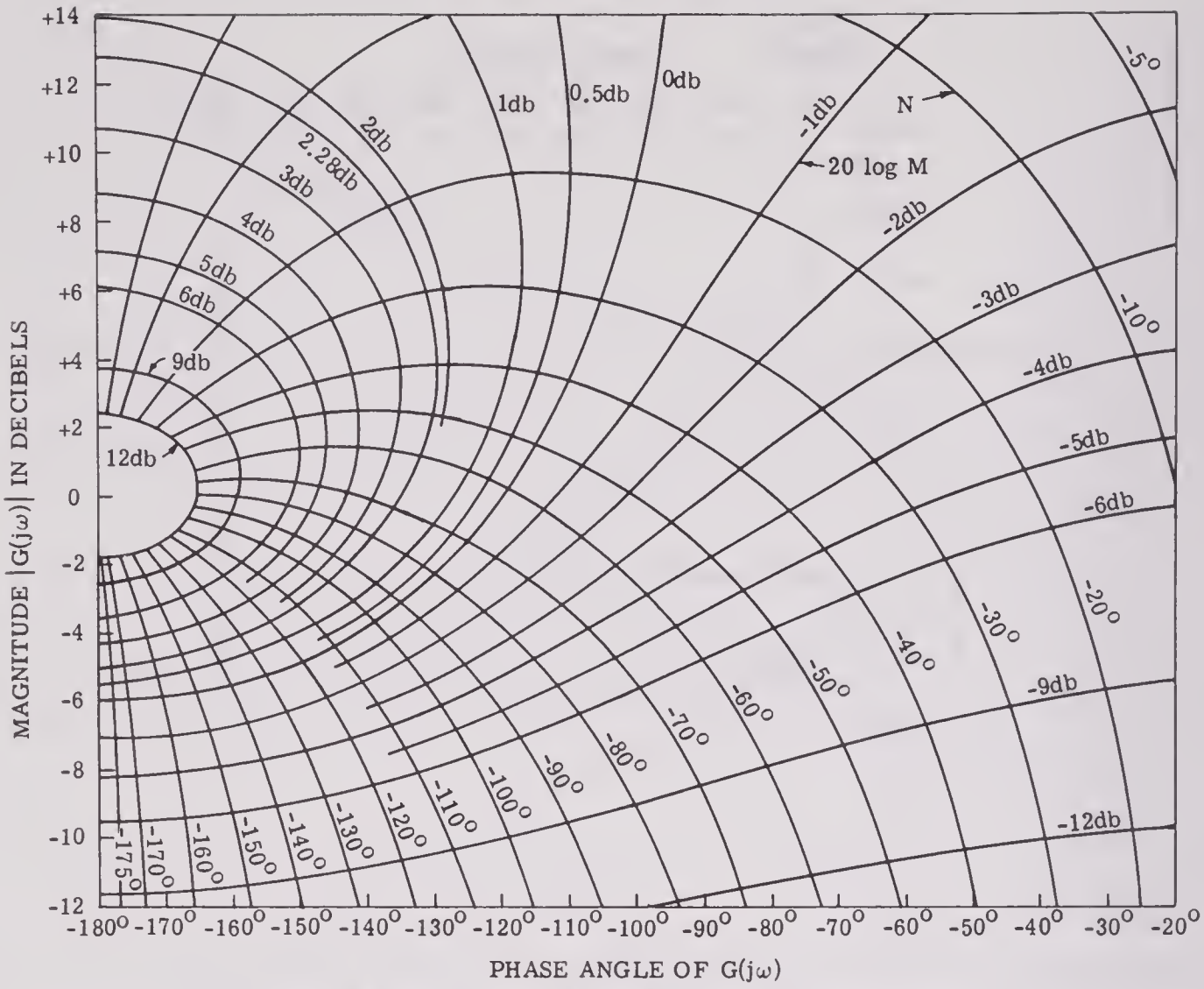


FIG. 17-23. M and N contours on log modulus plot (Nichols chart).

(f) The point at which these asymptote lines meet is determined by writing

$$G(s) = K \frac{s^m + a_1 s^{m-1} + a_2 s^{m-2} \ldots . a_m}{s^{m+n} + b_1 s^{m+n-1} + \ldots . am + n}$$

Then

$$s_1 = -\frac{b_1 - a_1}{n}$$

or

$$s_1 = \frac{\Sigma \text{poles} - \Sigma \text{zeros}}{(\text{number of finite poles}) - (\text{number of finite zeros})}$$

(g) Real-axis loci: Those portions of the real axis to the left of an odd number of critical frequencies (poles + zeros) are part of the root loci.

(h) Imaginary axis intersections: Apply the Routh test to $p(s) + q(s)$ (see Sec. 17.1.3.1).

(i) Angles of departure and arrival: At any point on the locus the sum of angles of vectors from poles must be $180° + 360n°$. The angle of departure can be chosen to satisfy this requirement.

(j) The point-α at which loci leave the real axis: Choose a point s_1 just off the axis (ϵ). The transition from $-\alpha$ to s_1 must yield no net change of angles from poles and zeros ($\theta_1 + \theta_2 + \theta_3 = 0$). Replace angles by their tangents and solve for α. If s_1 is above the real axis, conjugate complex roots will cause a net decrease in angle equal to

$$2\epsilon\gamma/(\beta^2 + \gamma^2)$$

where γ is the distance along the real axis from the complex poles (to the right of $-\alpha$) to $-\alpha$
β is the distance of the complex root from the real axis

(*k*) Sum of loci: If $p(s) + q(s)$ is of degree n and the coefficient of s^n is 1, the coefficient of s^{n-1} is the negative of the sum of the zeros.

(*l*) Product of loci: If $p(s)/q(s)$ contains a pole at the origin, the constant term of $p(s) + q(s)$ is directly proportional to K. Choose a point z on the locus. The gain at z is given by the value of K which makes $[p(s) + q(s)]/(s + z) = 0$. K at any point can also be determined by multiplying the lengths of the vectors from the open-loop poles and zeros.

ABSOLUTE STABILITY ANALYSIS. The system becomes unstable for values of K such that the roots are in the right-half plane.

STEADY-STATE BEHAVIOR. To determine steady-state behavior, determine the number of poles at the origin of the open-loop system. (See also Sec. 17.1.4.) If the number of poles is

(*a*) 0, the system has displacement error

(*b*) 1, the system has velocity error

(*c*) 2, the system has acceleration error

TRANSIENT BEHAVIOR. The contribution of a pole to system behavior is $Ae^{(\sigma+j\omega)t}$. The more negative σ, the less the contribution compared with roots closer to the $j\omega$ axis.

The frequency response function can be found as follows:

(*a*) Choose an ω on the $j\omega$ axis

(*b*) Let $G(s) = KK_s/(j\omega - r_1)(j\omega - r_2)\cdots$ where $j\omega - r_i$ is the distance from the chosen $j\omega$ to the ith root (for a given K).

(*c*) Graphical procedures are found in [2].

The time response can be found directly as follows:

(*a*) Write the open-loop transfer function as

$$KG(s) = \frac{K(a_1 s + b_1)(a_2 s + b_2)\cdots}{s^m(c_1 s + d_1)(c_2 s + d_2)\cdots}$$

$$= \frac{KK_s(s + 1/\tau_1)(s + 1/\tau_2)\cdots}{s^m(s + 1/\tau_a)(s + 1/\tau_b)\cdots}$$

where

$$K_s = K_1 K_2 K_3 \cdots \frac{\tau_1\tau_2\tau_3\cdots}{\tau_a\tau_b\tau_c\cdots}$$

(*b*) Take the inverse Laplace transform of

$$C(s) = \frac{KK_s R(s)}{(s - r_a)(s - r_b)\cdots}$$

where r_i is the location of the ith root in the closed-loop system.

(*c*) Graphical procedures are found in [2].

17.1.4. System Types and Performance

17.1.4.1. *Introduction.* Consider the system presented earlier in Fig. 17-17, where

$KG(s) = \dfrac{C(s)}{E(s)}$, the open-loop transfer function

$$G_c(s) = \frac{C(s)}{R(s)} = \frac{KG(s)}{1 + KG(s)}, \text{ the closed-loop transfer function}$$

$$\frac{E(s)}{R(s)} = \frac{1}{1 + KG(s)}, \text{ the error-response function}$$

A system is characterized by the open-loop transfer function as:

$$KG(s) = \frac{C(s)}{E(s)} = \frac{K(s\tau_1 + 1)(s\tau_2 + 1) \cdots}{s^N(s\tau_a + 1)(s\tau_b + 1) \cdots}$$

The system is classified

(*a*) Type 0 if $N = 0$

(*b*) Type 1 if $N = 1$

(*c*) Type 2 if $N = 2$

If a Bode plot is made for $C(s)$ for a given $R(s)$, the plot may be considered an approximation to an inverse time plot, i.e., $t \approx 1/\omega$.

Error coefficients are defined classically as follows:

(*a*) Position error coefficient: $K_p = \lim_{s \to 0} G(s)$

(*b*) Velocity error coefficient: $K_r = \lim_{s \to 0} sG(s)$

(*c*) Acceleration error coefficient: $K_a = \lim_{s \to 0} s^2G(s)$

Error coefficients can also be found from

$$\frac{E(s)}{R(s)} = \frac{1}{1 + KG(s)} = \frac{1}{1 + K_p} + \frac{s}{K_v} + \frac{s^2}{K_a}$$

Steady-state errors can be found for various inputs, since

(*a*) if $r(t) = h(t)$, $e_{ss} = (1 + K_p)^{-1}$

(*b*) if $r(t) = t$, $e_{ss} = (K_v)^{-1}$

(*c*) if $r(t) = t^2$, $e_{ss} = (K_a)^{-1}$

This is useful only for systems for which the poles of $G(s)$ are in the left-half s-plane.

17.1.4.2. *Type 0 System (Simple Speed Controllers, Feedback Amplifiers).* The open-loop transfer function of a Type 0 system is given by (see Fig. 17-24*a*)

$$KG(s) = \frac{K_p}{1 + \tau s}$$

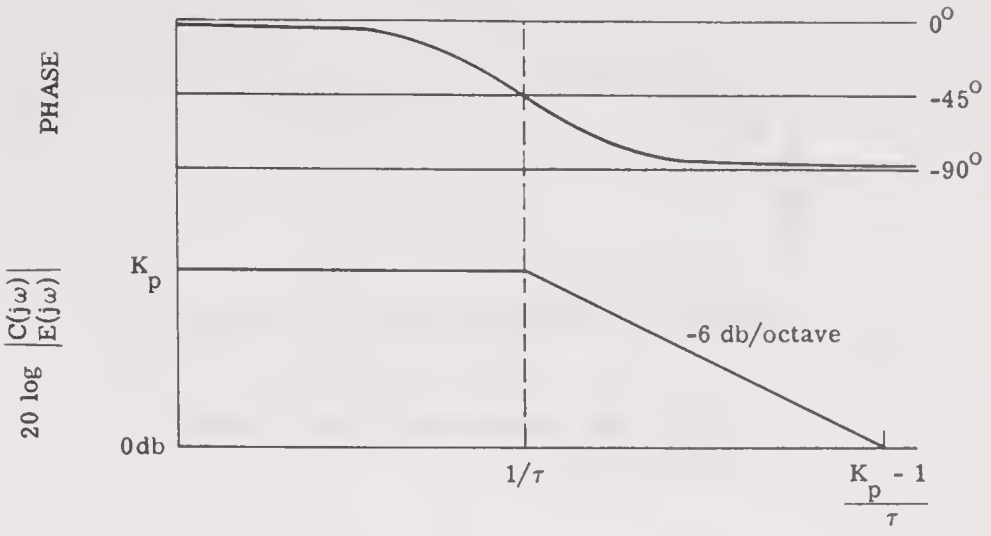
The closed-loop transfer function of a Type 0 system is given by

$$G_c(s) = \frac{K_p}{\tau s(1 + K_p)} \approx \frac{K_p}{1 + K_p}$$

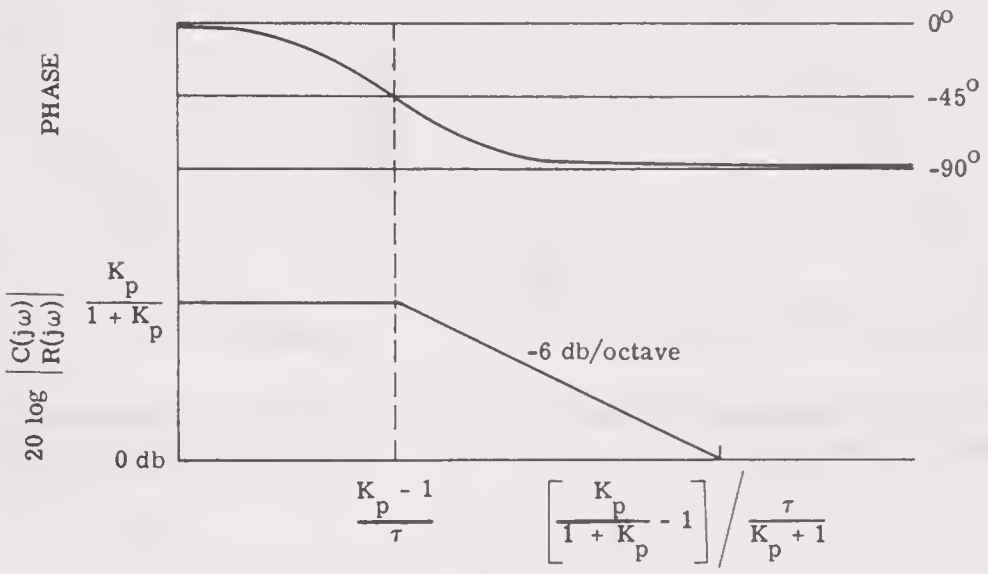
if $\tau << (1 + K_p)$.

The error (see Fig. 17-24*b*) is given by

$$\frac{E(s)}{R(s)} = \frac{1 + \tau s}{\tau s + (1 + K_p)} \approx \frac{1}{1 + K_p}$$



(a)



(b)

FIG. 17-24. Type 0 system Bode plot.
(*a*) Open loop. (*b*) Closed loop.

The response of the Type 0 system to a step input is described by the following equations (see Fig. 17-25).

$$r(t) = Ah(t) \qquad R(s) = \frac{A}{s}$$

$$C(s) = \frac{AK_p}{s(1 + K_p)} \qquad c(t) \approx \frac{K_p}{1 + K_p}$$

$$E(s) \approx \frac{A}{s(1 + K_p)} \qquad e(t) = \frac{A}{(1 + K_p)}$$

A Type 0 system develops a displacement error in response to a displacement input.

The response of the Type 0 system to ramp input is described by the following equations (see Fig. 17-26).

$$r(t) = \omega_i t \qquad R(t) = \frac{\omega_i}{s^2}$$

$$C(s) = \frac{K_p}{1 + K_p} \frac{\omega_i}{s^2} \qquad C(t) = \frac{K_p \omega_i t}{1 + K_p}$$

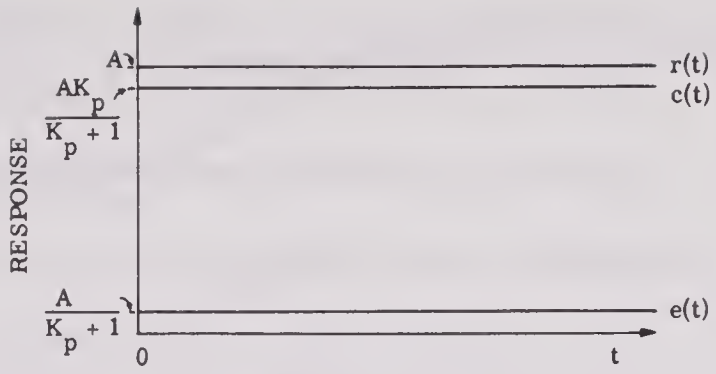


FIG. 17-25. Response of Type 0 system to step input.

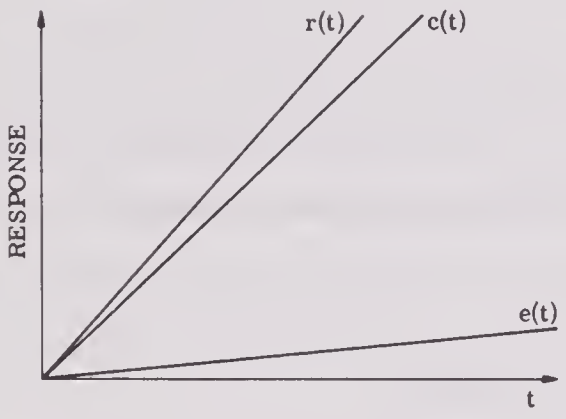


FIG. 17-26. Response of Type 0 system to ramp input.

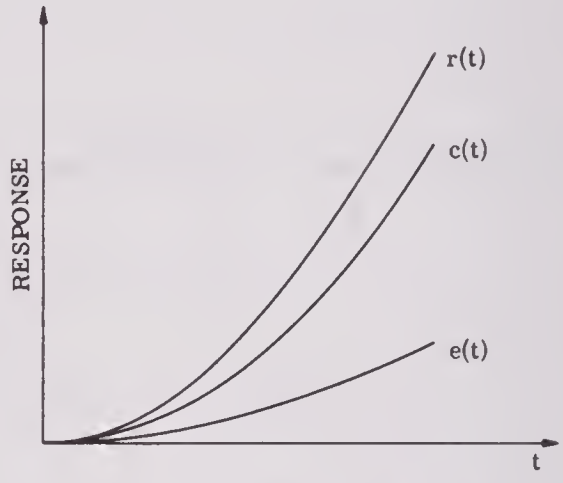


FIG. 17-27. Response of Type 0 system to parabolic input.

$$E(s) = \frac{1}{1 + K_p} \frac{\omega_i}{s^2} \qquad e(t) = \frac{\omega_i t}{1 + K_p}$$

A Type 0 system develops a velocity error in response to velocity input.

The response of the Type 0 system to parabolic input is described by the following equations (see Fig. 17-27).

$$r(t) = \alpha_i t^2 \qquad c(t) = \frac{\alpha_i}{s^3}$$

$$C(s) = \frac{K_p}{1 + K_p} \frac{\alpha_i}{s^3} \qquad c(t) = \frac{K_p \alpha t^2}{2(1 + K_p)}$$

$$E(s) = \frac{1}{1 + K_p} \frac{\alpha}{s^3} \qquad e(t) = \frac{\alpha t^2}{2(1 + K_p)}$$

A Type 0 system develops an acceleration error in response to an acceleration input.

17.1.4.3. *Type 1 System (Classical Motor-Driven Loops).* The open-loop transfer function of a Type 1 system is given by

$$\frac{C(s)}{E(s)} = \frac{K_v}{s(1 + \tau_m s)}$$

or, if $\tau_m << 1$ (see Fig. 17-28*a*),

$$\frac{c(s)}{E(s)} \approx \frac{K_v}{s}$$

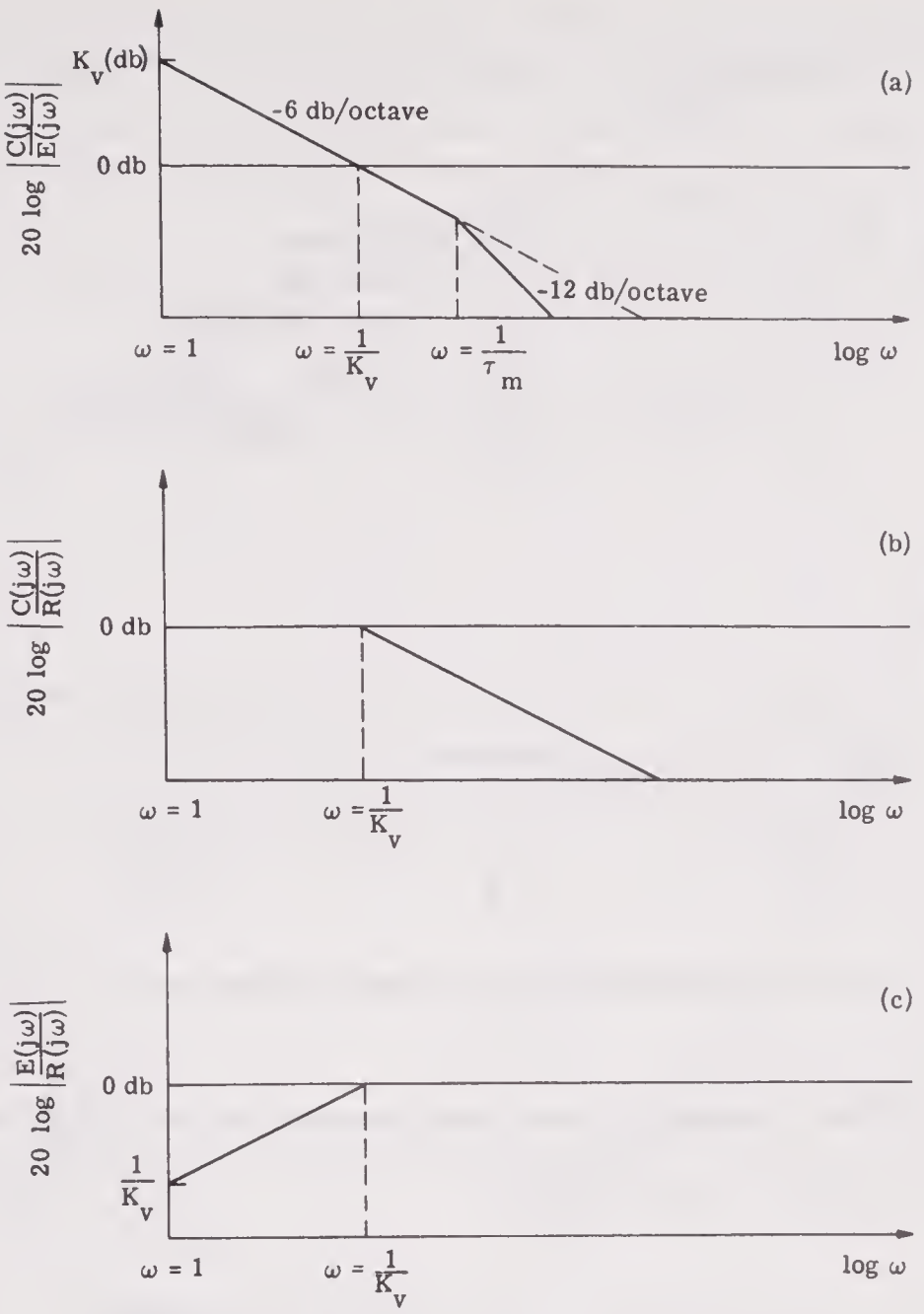


FIG. 17-28. Type 1 system Bode plot. (*a*) Open loop. (*b*) Closed loop. (*c*) Error response.

The closed-loop transfer function of a Type 1 system is given by (see Fig. 17-28*b*)

$$\frac{C(s)}{R(s)} = \frac{K_v}{K_v + s}$$

The error function is given by (see Fig. 17-28*c*)

$$\frac{E(s)}{R(s)} = \frac{s}{K_v + s} = \frac{(1/K_v)s}{1 + (s/K_v)}$$

The response of a Type 1 system step input is given by (see Fig. 17-29)

$$r(t) = Ah(t) \qquad R(s) = \frac{A}{s}$$

$$C(s) = \frac{A}{s[1 + s/K_v]} \qquad c(t) = A(1 - e^{-K_v t})$$

$$E(s) = \frac{1/K_v A}{(1 + s/K_v)} \qquad e(t) = Ae^{-K_v t}$$

A Type 1 system does not develop an error in response to a displacement input.

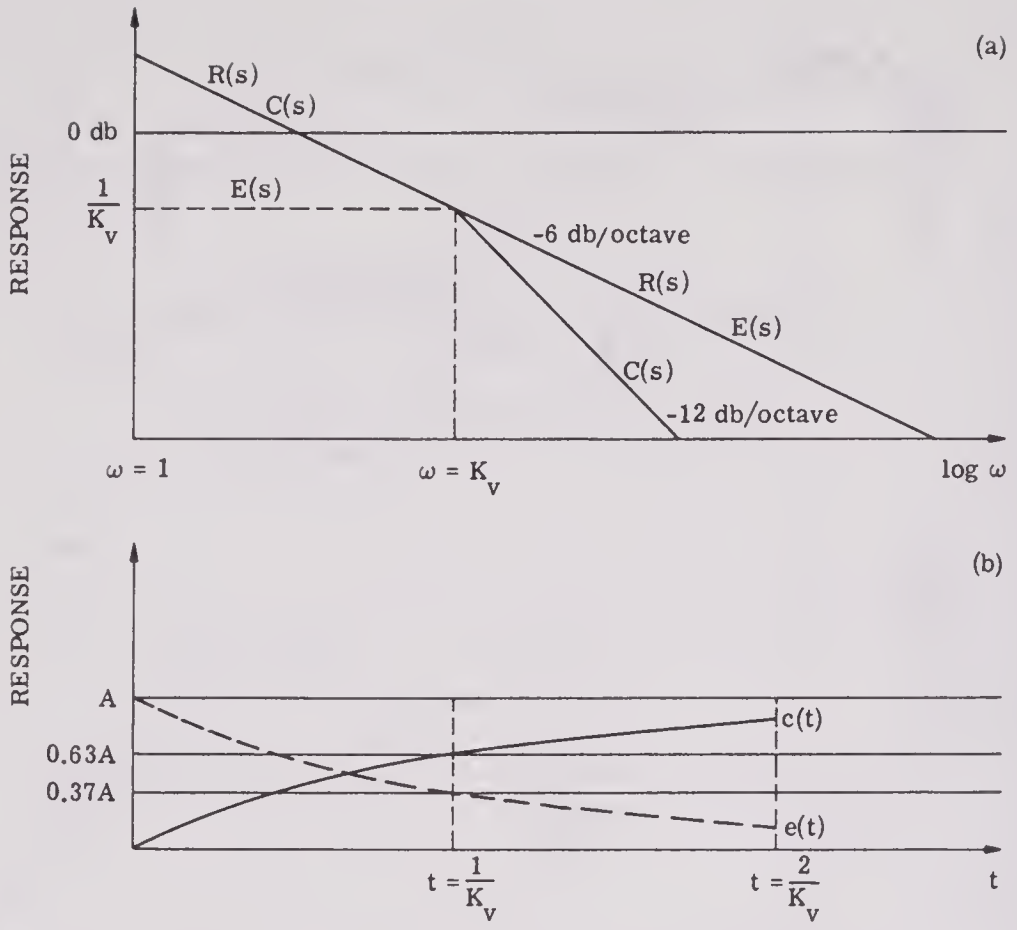


FIG. 17-29. Response of Type 1 system to step input.

The response of a Type 1 system to ramp input is given by (see Fig. 17-30)

$$r(t) = \omega_i t \qquad R(s) = \frac{\omega_i}{s^2}$$

$$C(s) = \frac{\omega}{s^2[1 + (s/K_v)]} \qquad c(t) = \omega t - \frac{\omega_i}{K_v}(1 - e^{-K_v t})$$

$$E(s) = \frac{\omega_i/K_v}{s(1 + s/K_v)} \qquad e(t) = \frac{\omega_i}{K_v}(1 - e^{-K_v t})$$

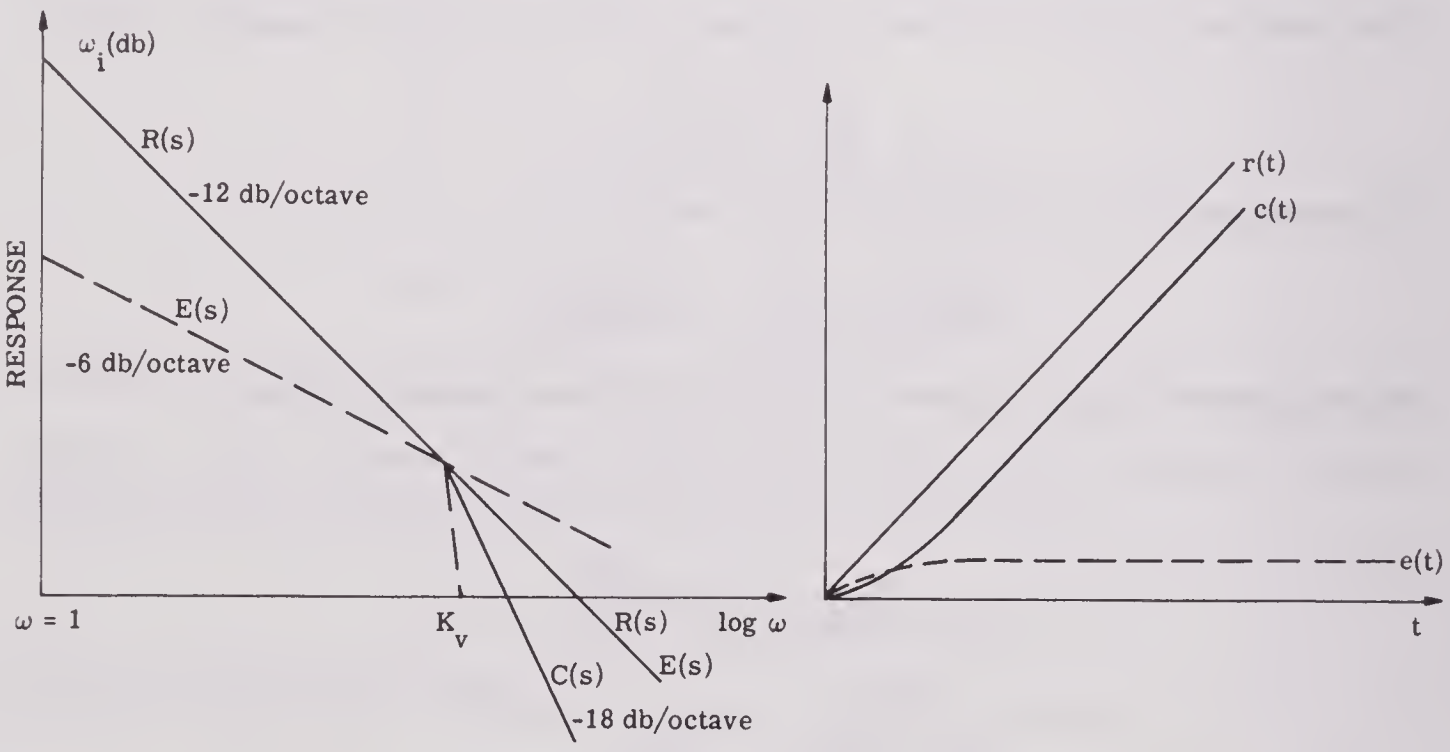


FIG. 17-30. Response of Type 1 system to ramp input.

A Type 1 system develops a displacement error ω_i/K_v in response to a ramp input $\omega_i t$. The response of a Type 1 system to parabolic input is given by (see Fig. 17-31)

$$r(t) = \frac{\alpha_i t^2}{2} \qquad R(s) = \frac{\alpha_i}{s^3}$$

$$C(s) = \frac{\alpha_i}{s^3[1 + (s/K_v)]} \qquad c(t) = \frac{\alpha_i t^2}{2} - \frac{\alpha_i}{K_v}\left(t - \frac{1 - e^{-K_v t}}{K_v}\right)$$

$$E(s) = \frac{\alpha_i/K_v}{s^2[1 + (s/K_v)]} \qquad e(t) = \frac{\alpha_i}{K_v}\left(t - \frac{1 - e^{-K_v t}}{K_v}\right)$$

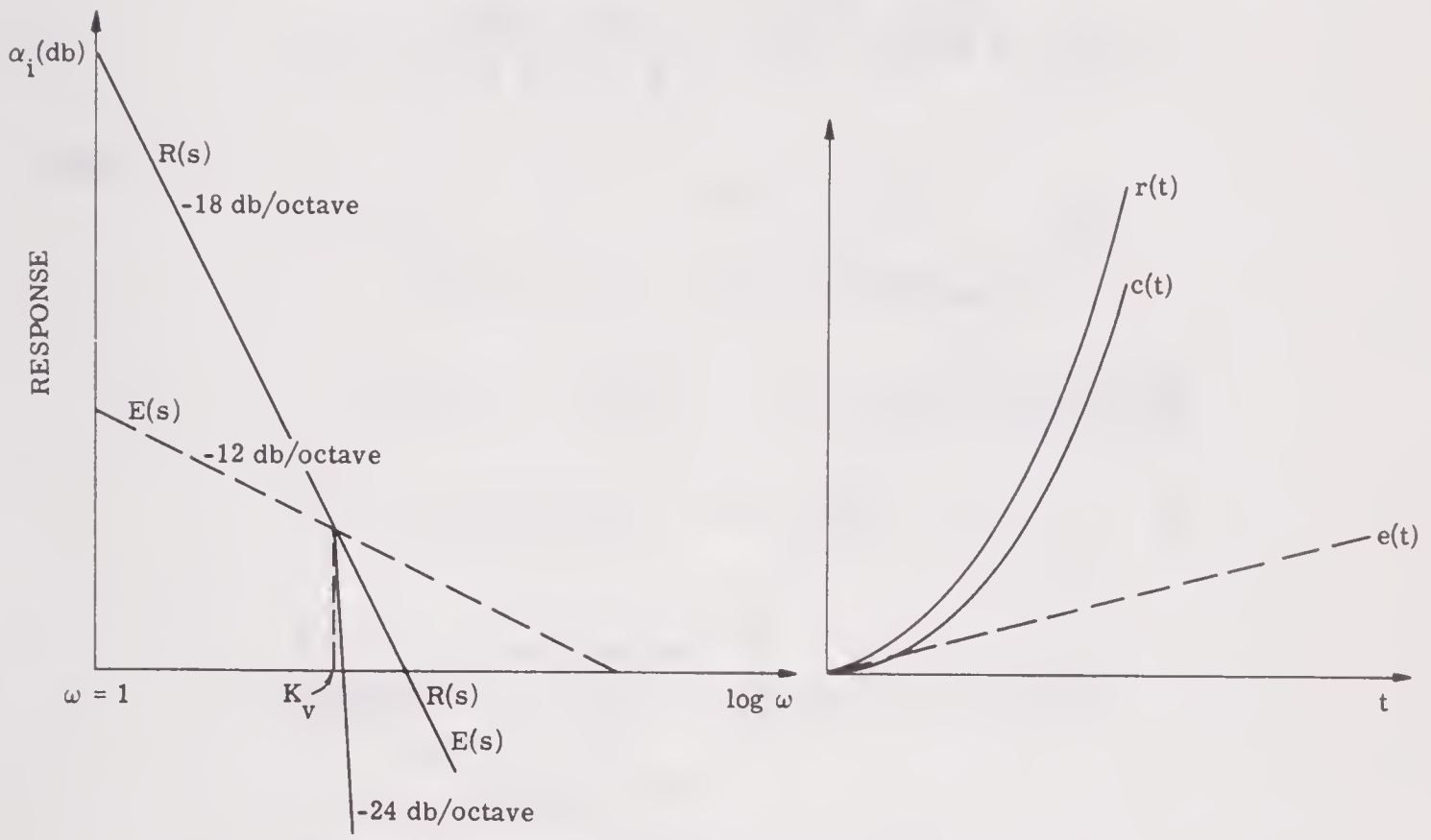


FIG. 17-31. Response of Type 1 system to parabolic input.

A Type 1 system develops a ramp error

$$\frac{\alpha}{K_v}\left(t - \frac{1}{K_v}\right)$$

in response to a parabolic input.

Notes on Behavior of Type 1 System

(*a*) The response is identical to the simple RC network.

(*b*) The time constant is $1/K_v$.

(*c*) The open-loop gain is unity at $\omega_c = K_v$; the loop is often referred to as closed to ω_c.

(*d*) To the first order, letting $t \approx 1/\omega$ (increasing time implies decreasing ω), at $t = 1/K_v (\omega = K_v)$, $c(t)$ merges smoothly with $r(t)$.

(*e*) The system gain at $t = 1/\omega$ is of the order of the frequency-response gain, and the error input signal is reduced by this gain at this time.

(*f*) The straight-line approximation to the system error shows that as time proceeds toward $t \approx 1/\omega = \infty$, the error is $r(t)$ until $t = 1/K_v (\omega = K_v)$; *i.e.* the system has not responded until this time.

(*g*) Note that $c(t) = K_v A e^{-K_v t} = K_v e(t)$; *i.e.,* in a Type 1 system the output rate is proportional to the error.

(*h*) The Type 1 system has an open-loop gain of K_v at $\omega = 1$, and a gain of unity at $\omega = K_v$.

(*i*) A Type 1 system reduces the resultant error by one order of t; *i.e.,* no error for displacement input; a displacement error for velocity input, and a velocity error for parabolic input.

17.1.4.4. *Type 2 System* (*Many Tracking Loops, Gearless Torqued Loops*). The open-loop transfer function of a Type 2 system is given by (see Fig. 17-32*a*)

$$\frac{C(s)}{E(s)} = \frac{K_a(1 + \tau s)}{s^2} \qquad (\tau << 1)$$

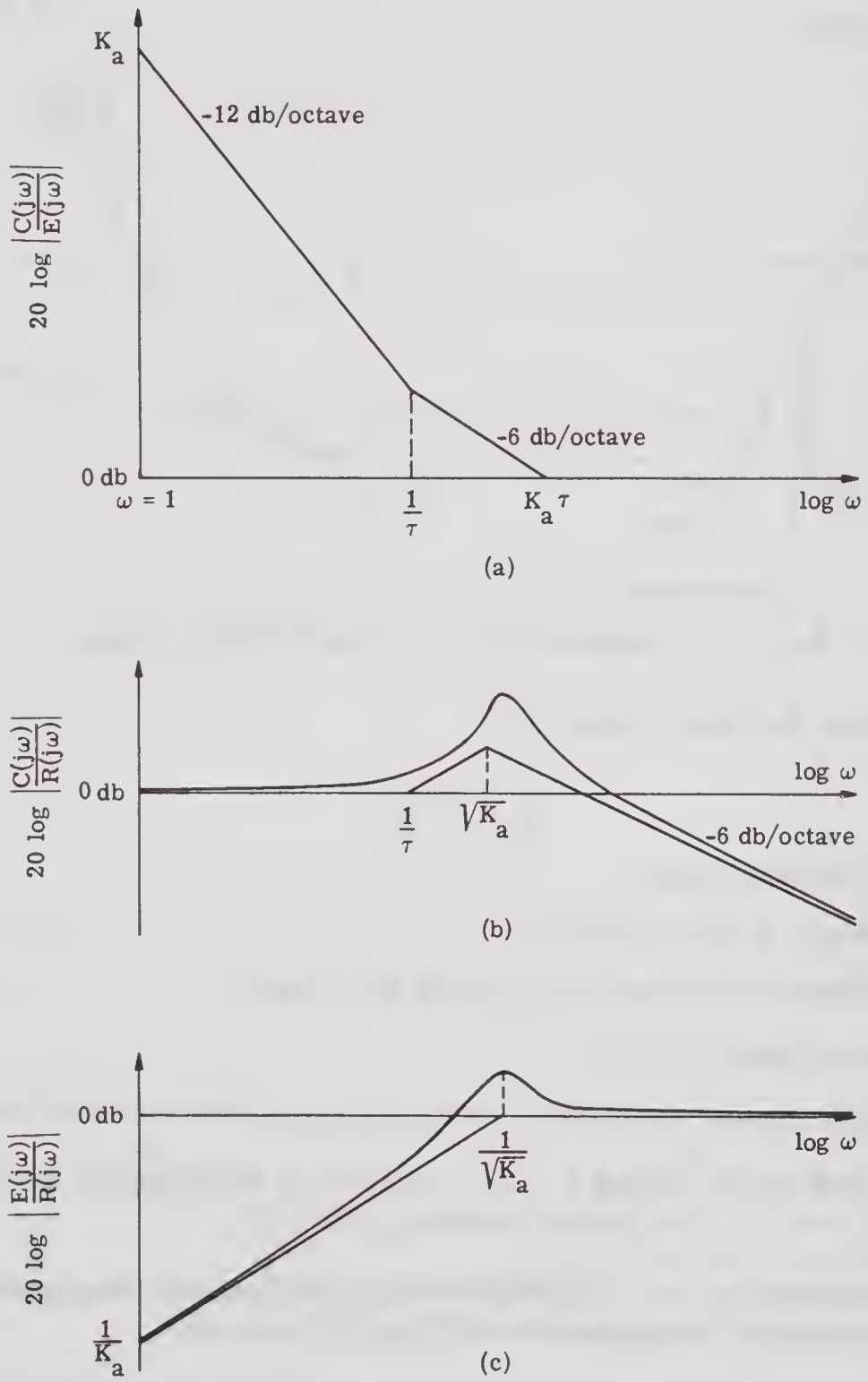


FIG. 17-32. Type 2 system Bode plot. (*a*) Open loop. (*b*) Closed loop. (*c*) Error response.

The closed-loop transfer function of a Type 2 system is given by (see Fig. 17-32*b*)

$$\frac{C(s)}{R(s)} = \frac{K_a(1 + \tau s)}{s^2 + K_a\tau s + K_a}$$

and (see Fig. 17-32*c*)

$$\frac{E(s)}{R(s)} = \frac{s^2}{s^2 + K_a\tau s + K_a}$$

The response of a Type 2 system to step input is given by (see Fig. 17-33)

$$r(t) = A \qquad R(s) = \frac{A}{s}$$

$$C(s) = \frac{AK_a(1 + \tau s)}{s(s^2 + K_a\tau s + K_a)} \qquad c(t) = A\left[1 - Ne^{\frac{-K_a\tau}{2}t}\right]$$

where

$$N = \begin{cases} 1 & [(K_a\tau)^2 - 4K_a] > 0 \\ \sin(\theta t + \psi) & [(K_a\tau)^2 - 4K_a] < 0 \end{cases}$$

$$\theta = \left[K_a\left(1 - \frac{K_a\tau^2}{4}\right)\right]^{1/2} \qquad \text{(frequency of oscillation)}$$

$$E(s) = \frac{As}{s^2 + K_a\tau s + K_a} \qquad e(t) = NA\, e^{-K_a\tau t/2}$$

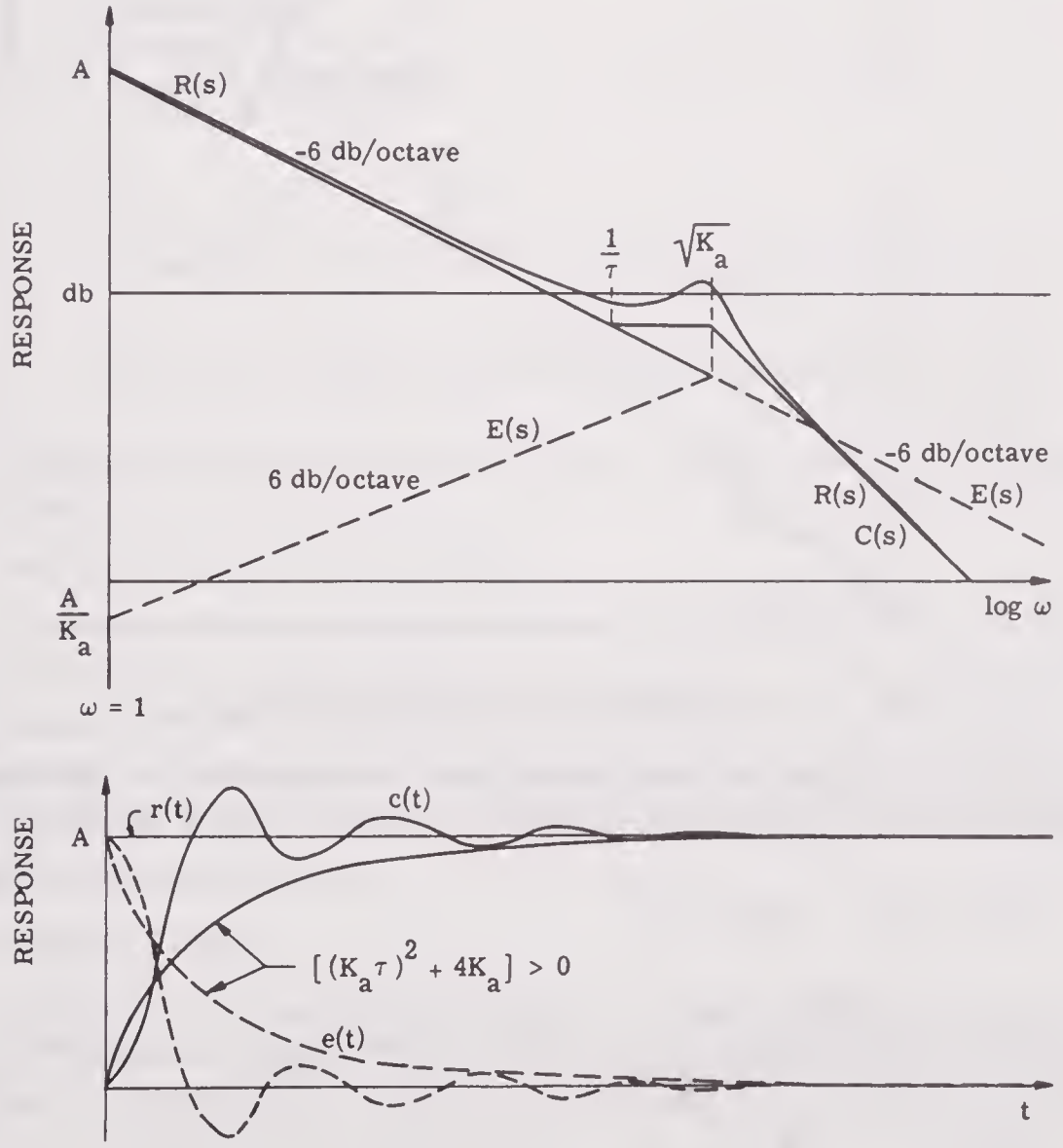


FIG. 17-33. Response of Type 2 system to step input.

The response of a Type 2 system to ramp input is given by (see Fig. 17-34)

$$r(t) = \omega_i t \qquad R(s) = \frac{\omega_i}{s^2}$$

$$C(s) = \frac{\omega_i K_a(1 + \tau s)}{s^2(s^2 + K_a\tau s + K_a)} \qquad c(t) = \omega_i(N\, e^{-K_a\tau t/2})$$

$$E(s) = \frac{\omega_i}{s^2 + K_a\tau s + K_a} \qquad e(t) = N\omega_j\, e^{-K_a\tau t/2}$$

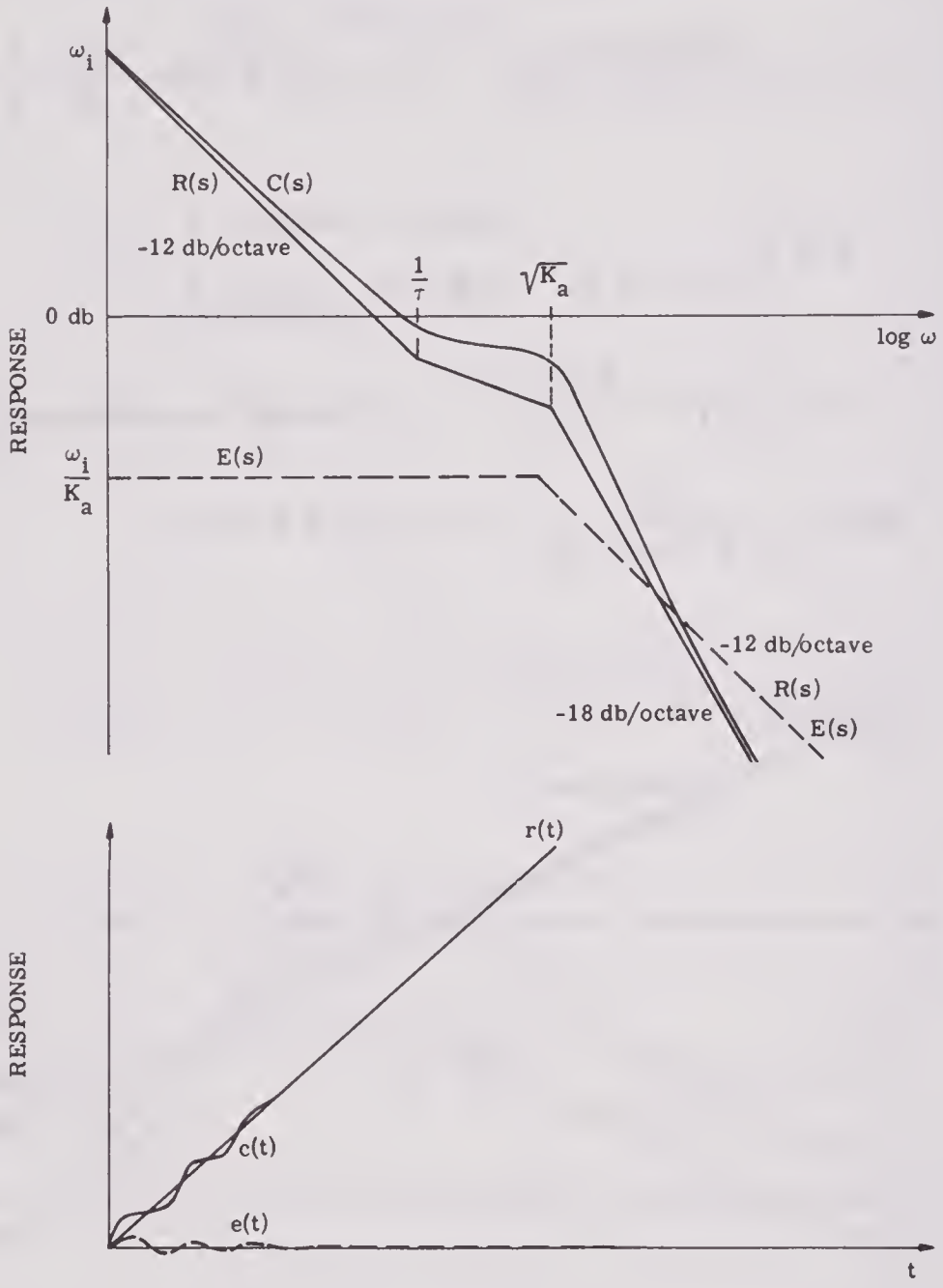


FIG. 17-34. Response of Type 2 system to ramp input.

A Type 2 system develops no steady-state error in response to a ramp input. The response of a Type 2 system to parabolic input is given by (see Fig. 17-35)

$$r(t) = \alpha t^2/2 \qquad R(s) = \frac{\alpha}{s^3}$$

$$\theta_0(s) = \frac{\alpha K_a(1 + \tau s)}{s^3(s^2 + K_a\tau s + K_a)} \qquad \theta_0(t) = \frac{1}{2}\,\alpha t^2 - \frac{\alpha}{K_a}\,(1 - N\, e^{-K_a\tau t/2})$$

$$\epsilon(s) = \frac{\alpha}{s(s^2 + K_a\tau s + K_a)} \qquad \epsilon(t) = \frac{\alpha}{K_a}\,(1 - N\, e^{-K_a\tau t/2})$$

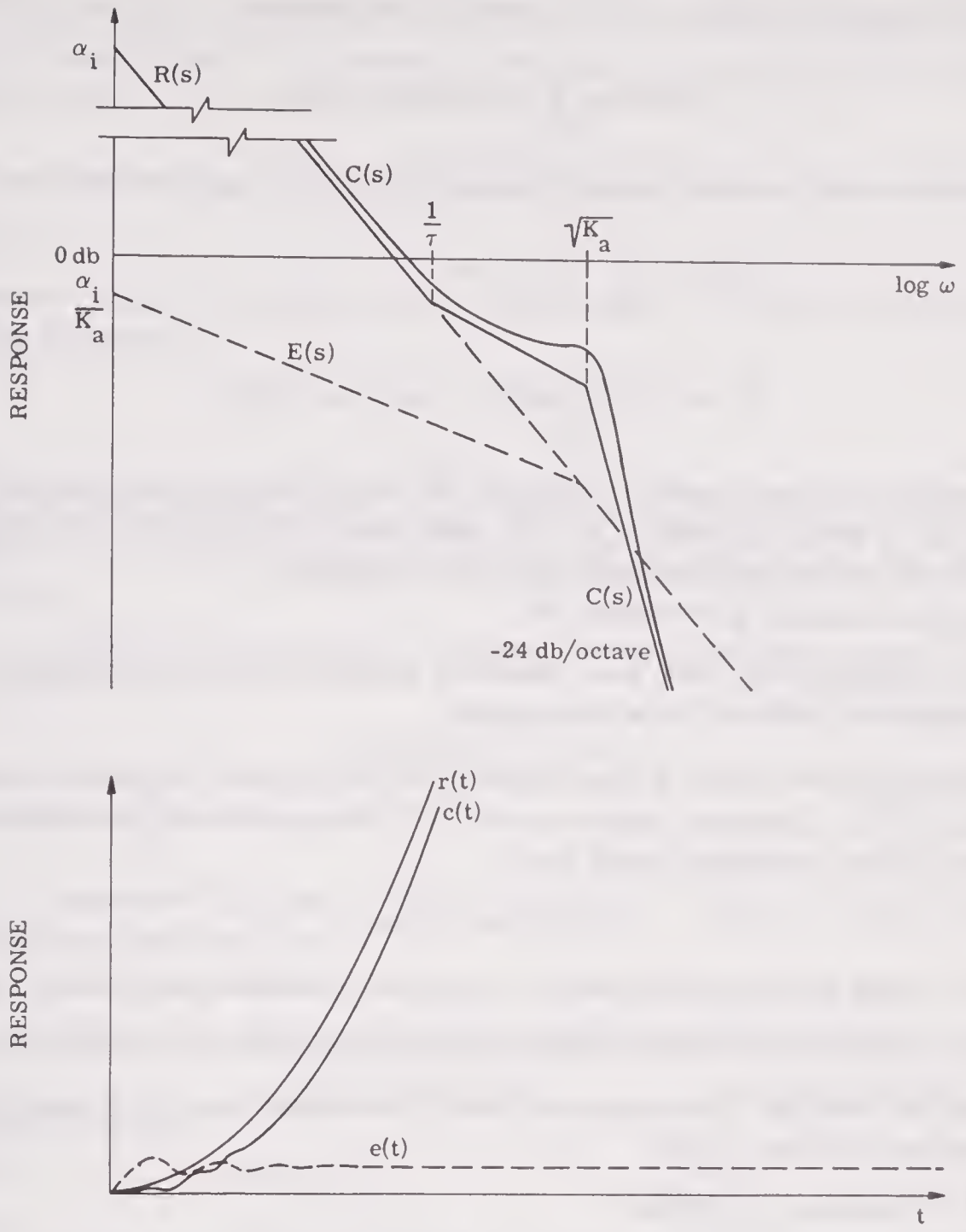


FIG. 17-35. Response of Type 2 system to parabolic input.

A Type 2 system develops a displacement error α/K_a in response to an acceleration (parabolic) input.

Notes on the Behavior of a Type 2 System

If $[(K_a\tau)^2 - 4K_a] < 0$, the response is a sharp oscillatory rise with overshoot. The transient decays with a time constant of $1/2\omega_c$ where ω_c is the crossover frequency. The initial response of this system is faster than that of a Type 1 system of the same closed-loop frequency (see Sec. 17.1.4.3), but the transient decay time is longer. Again, the time domain performance can be *crudely* approximated from the plot ($t \approx 1/\omega$).

17.2. Sampled-Data Systems [3]

17.2.1. Basic Definitions

17.2.1.1. *Sampling.* The sampler modulates a series of periodic narrow pulses with the incoming, continuous signal. It is assumed, in general, that the pulses have zero width. These can be represented as $\delta(t - nT)$, for which the Laplace transform is

$$\Delta(s) = \frac{1}{(1 - e^{-Ts})}$$

where T is the sampling period. If the input to the sampler is $r(t)$, the output is

$$r^*(t) = \sum_{n=0}^{\infty} r(nT)\delta(t - nT)$$

The transform relationships between the input and output can be written two ways:

$$R^*(s) = \sum_{n=0}^{\infty} r(nT)e^{-nTs}$$

$$R^*(s) = \frac{1}{T} \sum_{n=\infty}^{\infty} R(s + jn\omega_s) \quad \omega_s = 1/T$$

The first equation is most useful in analysis; the second indicates the modulation characteristic, *i.e.*, a series of carriers at $n\omega_s$, each modulated by $r(t)$. The frequencies around $n\omega_s$ are called complementary signals or sidebands.

Two important sampling theorems are:

(a) If the sampling rate is at least twice the highest frequency component, a signal can be transmitted without loss of information.

(b) The mean square value of the samples $\overline{f^2(kT)}$ is equal to the mean square value of the signal $\overline{f^2(t)}$. However, this is true only if the signal contains neither any frequencies $n\omega_s/2$ nor components such that

$$|\omega_1 \pm \omega_2| = n\omega_s$$

17.2.1.2. *Types of Sampled Systems.* Two types of sampled systems are as follows:

(a) Error sampled: The sampler follows immediately after the error-sensing device.

(b) Nonerror sampled: This system can usually be reduced to an error-sampled system by applying rules of Sec. 17.2.2.

17.2.1.3. *Clamping or Holding*

INTRODUCTION. In general, clamping or holding circuits act as phase-lag circuits; the higher-order the hold, the more lag. They are used to remove complementary frequencies caused by the sampling process, convert the samples into approximately continuous form, supply energy for the actuation of the output device, and provide a finite drive for the continuous portion of the system.

TYPES OF HOLDING CIRCUITS

Zero-Order or Boxcar

This circuit holds the measured value for part or all of the period between pulses:

$$c_n(t) = c(nT)$$

A full clamp holds the signal value until the next sample pulse. This produces less distortion but delays the signal by $T/2$. The transfer function is given by

$$G_{h0}(s) = (1 - e^{-Ts})/s$$

First-Order

This circuit approximates the signal between pulses as a first-order polynominal:

$$c_n(t) = c(nT) + c'(nT)(t - nT)$$

i.e., the slope is based on previously measured values, hence little error is generated for constant slope input. The necessity of obtaining an estimate of the slope causes an additional delay. The transfer function of this holding circuit is

$$G_{h1} = \frac{1 + Ts}{T}\left[\frac{1 - e^{-Ts}}{s}\right]^2$$

*k*th *Order*

The general *k*th-order holding circuit approximates the signal between pulses by a *k*th order polynomial:

$$c_n(t) = c(nT) + c'(nT)(t - nT) + \cdots$$

$$c^k(nT)(t - nT)^k$$

The result of such approximation is more effective smoothing, but a greater delay is encountered. This system is usually detrimental to stability, as well as complex and expensive.

17.2.1.4. *Weighting Sequence* $g_n(t)$. Output of a sampled-data system for unit impulse input (at $t = 0T$)

$$g_n(t) = g(nT) = g(t)\big|_{t=nT}$$

Note that $g^*(t) = \sum_{n=0}^{\infty} g_n(t)\delta(t - nT)$

17.2.1.5. *Equivalent Continuous Data Functions* [3]. Table 17-1 gives some equivalent continuous data functions.

TABLE 17-1. EQUIVALENT CONTINUOUS DATA FUNCTIONS [3]

		Continuous-Data System	*Pulsed-Data System*
1.	Input and output:		
	$\delta(t)$	$y(t) = w(t)$	$y_n = w_n$
	$u(t)$	$y(t) = \int_0^t w(\tau)\, d\tau$	$y_n = \sum_{k=0}^{n} w_k$
	$x(t)$	$y(t) = \int_0^t w(\tau)x(t - \tau)\, d\tau$	$y_n = \sum_{k=0}^{n} w_k\, x_{n-k}$
		$y(t) = \int_0^t x(\tau)w(t - \tau)\, d\tau$	$y_n = \sum_{k=0}^{n} x_k\, w_{n-k}$
2.	Frequency functions	$Y(j\omega) = A(\omega) + jB(\omega)$	$Y^*(j\omega) = A^*(\omega) + jB^*(\omega)$
		$A(\omega) = \int_0^\infty y(t) \cos \omega t\, dt$	$A^*(\omega) = \sum_{n=0}^{\infty} y(nT) \cos nT\omega$
		$B(\omega) = -\int_0^\infty y(t) \sin \omega t\, dt$	$B^*(\omega) = -\sum_{n=0}^{\infty} y(nT) \sin nT\omega$
3.	Time functions	$y(t) = \frac{1}{2\pi}\int_0^\infty A(\omega) \cos t\omega\, d\omega$	$y(nT) = \frac{2}{\pi}\int_0^\pi A^*(\theta) \cos n\theta\, d\theta$
			$y(nT) = -\frac{2}{\pi}\int_0^\pi B^*(\theta) \sin n\theta\, d\theta$
4.	Stability condition	$\int_0^\infty \lvert w(t)\rvert\, dt < \infty$	$\sum_{n=0}^{\infty} \lvert w_n\rvert < \infty$
5.	Described by	Differential equations	Difference equations

17.2.1.6. *z-Transforms.* The system considered is shown in Fig. 17-36. z-transforms permit the relationship between $R(s)$ and $C(s)$ to be written as

$$\frac{R(z)}{C(z)} = G(z)$$

where $z = e^{Ts}$ and $G(z)$ is the pulsed transfer function or z-transfer function for the system.

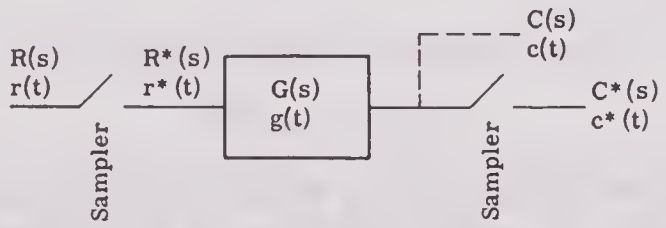


FIG. 17-36. Basic sampled data system.

Some basic relationships are:

$$R(z) = R\left(\frac{1}{T} \ln z\right) = \sum_{n=0}^{\infty} r(n,t)z^{-n}$$

$$R(z) = z\{r(t)\} = L\{r^*(t)\}\Big|_{s=\frac{1}{T}\ln z = z\{R(s)\}}$$

$$C(z) = z\{G(s)R^*(s)\}$$

A list of z-transforms is given in Table 17-2.

Some theorems relating to z-transforms follow:

(a) They are linear

(b) $z\{g(t - nT)\} = z^{-n}G(z)$

$z\{g(t + nT)\} = z^{+n}G(z)$

if $n > 0$ and $g(kT) = 0$ for $0 < k < (n - 1)$.

(c) $z\{e^{at}g(t)\} = G(e^{-aT}z)$

$z\{e^{-at}g(t)\} = G(e^{aT}z)$

(d) $\lim_{z \to 1}\left\{\frac{z-1}{z} G(z)\right\} = \lim_{t \to \infty}\{g(t)\}$

(e) $\lim_{z \to \infty}\{G(z)\} - \lim_{t \to 0}\{g(t)\}$

The following lists some general properties of z-transforms.

(a) They are periodic, with period ω_s.

(b) $G(z)$ is real at $z = 1(\omega = 0)$ and $z = 1(\omega = n\omega_s/2)$.

(c) The poles of $G(z)$ are $z_k = e^{Ts}k$.

(d) Shifting the zeros of $G(s)$ changes the gain constants in $G(z)$.

(e) If $G(s)$ has simple poles only and no pole at the origin, $G(z)$ describes a closed curve as z traverses the unit circle of the z plane.

(f) If $G(s)$ has a pole at the origin, $G(z)$ closes clockwise at infinity; *i.e.,* there is a pole at $z = 1$.

TABLE 17-2. z TRANSFORMS [3]

	$G(s)$	$g(t)$	$G(z)$	$g(nT)$
1.	1	$\delta(t)$	1 or z^{-0}	$\delta(nT)$
2.	ϵ^{-kTs}	$\delta(t-kT)$	z^{-k}	$\delta(nT-kT)$
3.	$\dfrac{1}{s-\frac{1}{T}\ln a}$	$a^{t/T}$	$\dfrac{z}{z-a}$	a^n
4.	$\dfrac{1}{s}$	$u(t)$	$\dfrac{z}{z-1}$	$u(nT)$ or 1
5.	$\dfrac{1}{s^2}$	t	$\dfrac{Tz}{(z-1)^2}$	nT
6.	$\dfrac{1}{s^3}$	$\dfrac{1}{2!}t^2$	$\dfrac{T^2}{2}\dfrac{z(z+1)}{(z-1)^3}$	$\dfrac{1}{2}(nT)^2$
7.	$\dfrac{1}{s^k}$	$\dfrac{1}{(k-1)!}t^{k-1}$	$\lim\limits_{a\to 0}\dfrac{(-1)^{k-1}}{(k-1)!}\dfrac{\partial^{k-1}}{\partial a^{k-1}}\left(\dfrac{z}{z-\epsilon^{-aT}}\right)$	$(-1)^{k-1}\lim\limits_{a\to 0}\dfrac{\partial^{k-1}}{\partial a^{k-1}}(\epsilon^{-aTn})$
8.	$\dfrac{1}{s+a}$	ϵ^{-at}	$\dfrac{z}{z-\epsilon^{-aT}}$	ϵ^{-anT}
9.	$\dfrac{1}{(s+a)^2}$	$t\epsilon^{-at}$	$-\dfrac{Tz\epsilon^{-aT}}{(z-\epsilon^{-aT})^2}$	$(nT)\epsilon^{-anT}$
10.	$\dfrac{1}{(s+a)^{k+1}}$	$\dfrac{t^k}{k!}\epsilon^{-at}$	$(-1)^k\dfrac{1}{k!}\dfrac{\partial^k}{\partial a^k}\left(\dfrac{z}{z-\epsilon^{-aT}}\right)$	$(-1)^k\dfrac{1}{k!}\dfrac{\partial^k}{\partial a^k}(\epsilon^{-anT})$
11.	$\dfrac{a}{s(s+a)}$	$1-\epsilon^{-at}$	$\dfrac{(1-\epsilon^{-aT})z}{(z-1)(z-\epsilon^{-aT})}$	$1-\epsilon^{-anT}$
12.	$\dfrac{a}{s^2(s+a)}$	$t-\dfrac{1-\epsilon^{-at}}{a}$	$\dfrac{Tz}{(z-1)^2}-\dfrac{(1-\epsilon^{-aT})z}{a(z-1)(z-\epsilon^{-aT})}$	$nT-\dfrac{1-\epsilon^{-anT}}{a}$
13.	$\dfrac{\omega_0}{s^2+\omega_0^2}$	$\sin\omega_0 t$	$\dfrac{z\sin\omega_0 T}{z^2-2z\cos\omega_0 T+1}$	$\sin n\,\omega_0 T$
14.	$\dfrac{s}{s^2+\omega_0^2}$	$\cos\omega_0 t$	$\dfrac{z(z-\cos\omega_0 T)}{z^2-2z\cos\omega_0 T+1}$	$\cos n\,\omega_0 T$
15.	$\dfrac{\omega_0^2}{s(s^2+\omega_0^2)}$	$1-\cos\omega_0 t$	$\dfrac{z(1-\cos\omega_0 T)(z+1)}{(z-1)(z^2-2z\cos\omega_0 T+1)}$	$1-\cos n\,\omega_0 T$
16.	$\dfrac{\omega_0}{(s+a)^2+\omega_0^2}$	$\epsilon^{-at}\sin\omega_0 t$	$\dfrac{z\epsilon^{-aT}\sin\omega_0 T}{z^2-2\epsilon^{-aT}z\cos\omega_0 T+\epsilon^{-2aT}}$	$\epsilon^{-anT}\sin n\,\omega_0 T$

The z-transform maps the left half plane into the inside of a unit circle, and the negative real axis into the positive real axis.

17.2.1.7. *Inverse z-Transforms.* The basic method for finding the inverse transform is to apply

$$g(nT) = \frac{1}{2\pi j} \oint_{\Gamma} G(z) z^{n-1}\, dz$$

where Γ is the path enclosing all singularities of $G(z)z^{n-1}$.

The inverse transform can often be evaluated more easily by using

$$g(nT) = \sum_{z_k} (\text{residues of } G(z)z^{n-1} \text{ at } z_k)$$

Another method is to perform a partial fraction expansion; then a table of transforms can be used to yield an approximate output.

Still another method is the use of a power series expansion (long division):

$$G(z) = \frac{b_m z^m + b_{m-1} z^{m-1} \cdots b_0}{C_n z^n + C_{n-1} z^{n-1} \cdots C_0} \qquad n \geq m$$

Division Yields

$$G(z) = a_0 + a_1 z^{-1} + a_2 z^{-2} t \cdots$$

where the a_i are the values of $g^*(t)$ at the ith sampling period.

17.2.2. Determination of Transfer Functions

17.2.2.1. *Transfer Function in s-Plane.* For the open-loop system of Fig. 17-36, the transfer function is determined as the ratio of input and output Laplace transforms, *i.e.*,

$$G(s) = \frac{C(s)}{R^*(s)}$$

$$G^*(s) = \frac{C^*(s)}{R^*(s)}$$

For the closed-loop system (Fig. 17-37) the following can be easily derived:

$$E(s) = R(s) - B(s)$$

$$E^*(s) = R^*(s) - B^*(s)$$

$$B(s) = G(s)H(s) = GH(s)$$

$$E^*(s) = R^*(s) - GH^*(s)$$

$$C(s) = E^*(s)\, G(s)$$

$$\frac{E^*(s)}{R^*(s)} = \frac{1}{1 + GH^*(s)}$$

$$\frac{C(s)}{R^*(s)} = \frac{G(s)}{1 + GH^*(s)}$$

$$GH^*(s) = \frac{1}{T} \sum_{n=-\infty}^{\infty} GH(s + jn\omega_s)$$

17.2.2.2. *Transfer Function in z-Plane.* Derivation of the transfer function from $G(s)$ can be accomplished by taking the inverse transform to obtain $g(t)$. Then, substituting $t = nT$ will permit the application of

$$G(z) = \sum_{n=0}^{\infty} g(nT) z^{-n}$$

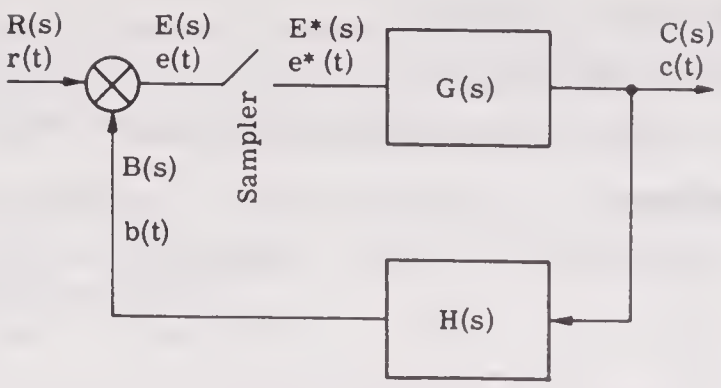


FIG. 17-37. Sampled data feedback system.

or, for complicated $G(s)$, partial fractions can be used. Some z-transforms are given in Table 17-2.

In the block diagram in Fig. 17-38, quantities or blocks not separated by a sampler are combined, and the z-transform of the combination taken.

$$C(z) = G_2(z)G_1R(z)$$

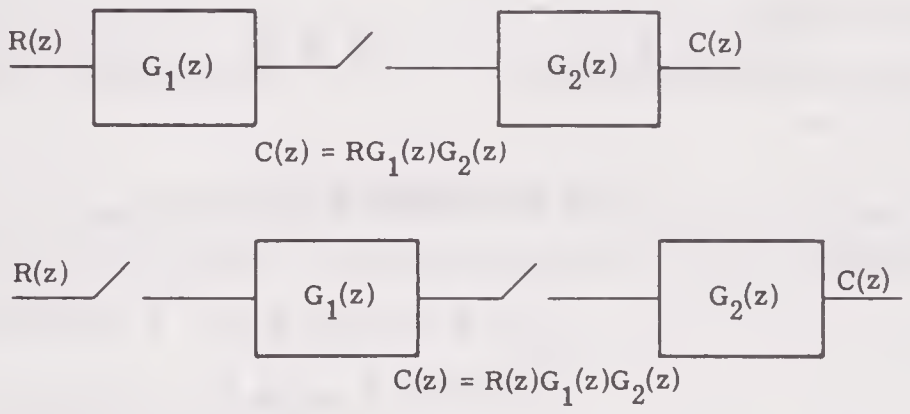


FIG. 17-38. Block diagram transfer-function derivation.

Quantities or blocks separated by a sample are transformed and then combined:

$$C(z) = G_1(z)G_2(z)R(z)$$

The rules above can be applied for feedback systems, or the system output can be written in terms of the Laplace transform. Then let $G(s) \rightarrow G(z)$; $G(s)H(s) \rightarrow GH(z)$.

17.2.3. Methods of Analyzing Sampled-Data Systems

17.2.3.1. *Directly from Differential Equation*

REPRESENTATION OF THE SYSTEM. In Fig. 17-37, let

$$A(z) = GH(z)$$

The characteristic equation $F(z)$ is then the numerator of $1 + A(z)$.

ABSOLUTE STABILITY ANALYSIS. If $F(z)$ is a quadratic polynomial with real coefficients, and the coefficient of z^2 is 1, the system is stable when all three of the following conditions hold:

$$|F(0)| < 1$$

$$F(1) > 0$$

$$F(-1) > 0$$

THE SCHUR-COHN STABILITY CRITERION

Write the characteristic equation as

$$F(z) = a_0 + a_1 z + a_2 z^2 \cdots a_n z^n = 0$$

Then form the determinant

$$\Delta_k = \begin{vmatrix} a_0 & 0 & 0 & \cdots & 0 & a_n & a_{n-1} & \cdots & a_{n-k+1} \\ a_1 & a_0 & 0 & \cdots & 0 & 0 & a_n & \cdots & a_{n-k+2} \\ & \cdots & a_0 & \cdots & & & & & \vdots \\ & & \cdots & & & & & & \vdots \\ a_{k-1} & a_{k-2} & a_{k-3} & \cdots & a_0 & 0 & 0 & \cdots & a_n \\ \overline{a}_n & 0 & 0 & \cdots & 0 & \overline{a}_0 & \overline{a}_1 & \cdots & \overline{a}_{k-1} \\ \overline{a}_{n-1} & \overline{a}_n & 0 & \cdots & 0 & 0 & \overline{a}_0 & \cdots & \overline{a}_{k-2} \\ & \cdots & \overline{a}_n & \cdots & & & & & \vdots \\ & & \cdots & & & & & & \vdots \\ \overline{a}_{n-k+1} & \overline{a}_{n-k+2} & \cdots & & \overline{a}_n & 0 & 0 & \cdots & \overline{a}_0 \end{vmatrix}$$

where $k = 1, 2, \cdots n$

$\overline{a}_n$ is the conjugate of a_n

Example for $k = 3$:

$$\Delta_3 = \begin{vmatrix} a_0 & 0 & 0 & a_3 & a_2 & a_1 \\ a_1 & a_0 & 0 & 0 & a_3 & a_2 \\ a_2 & a_1 & a_0 & 0 & 0 & a_3 \\ \overline{a}_3 & 0 & 0 & \overline{a}_0 & \overline{a}_1 & \overline{a}_2 \\ \overline{a}_2 & \overline{a}_3 & 0 & 0 & \overline{a}_0 & \overline{a}_1 \\ \overline{a}_1 & \overline{a}_2 & \overline{a}_3 & 0 & 0 & \overline{a}_0 \end{vmatrix}$$

A system is stable if $\Delta_k < 0$ for k odd

$\Delta_k > 0$ for k even

17.2.3.2. *Nyquist Method*

REPRESENTATION OF THE SYSTEM. Construct a polar plot of $G(z)$ as z goes once around the unit circle.

ABSOLUTE STABILITY ANALYSIS. The number of encirclements Ω of the critical point (-1 or $-1/k$) indicates stability, since if $\Omega = n - p$ where $p =$ number of poles of $G(z)$ and $n =$ number of zeros of $G(z)$, the system is stable.

TRANSIENT RESPONSE ANALYSIS. The transient behavior of a sampled-data system can be determined in a manner similar to that used in continuous systems.

17.2.3.3. *Bilinear Transformation Analysis.* Make the substitution in $G(z)$

$$z = \frac{1 + W}{1 - W}$$

This yields an expression which is the ratio of two polynomials of the same order in W. Continuous data system analysis techniques (such as Routh, Nyquist, Bode, and

root locus) can then be applied to determine stability, relative stability, etc. The open-loop transfer function in W of a sampled-data system is nonminimum phase; both gain and phase must be sketched in applying the Bode method.

17.2.3.4. *Root Locus Method.* The construction of root loci is carried out as for continuous systems, but in the z-plane. Table 17-3 indicates the expected behavior of a system from the pole location.

TABLE 17-3. EFFECT OF POLE AND ZERO LOCATION ON BEHAVIOR OF SAMPLED-DATA SYSTEM [3]

Location of Closed-Loop Poles	*Mode of Transient Behavior*
Outside the unit circle	Unstable operation
Inside the unit circle	Stable operation
(a) Real pole in right half	(a) Decaying sequence
(b) Real pole in left half	(b) Alternating decaying sequence
(c) Conjugate complex poles	(c) Damped oscillatory sequence

17.2.4. Types of Sampled-Data Systems. In general, the characteristics of continuous-data systems carry over to these sampled-data systems (see Sec. 17.1.4).

Type 0: *No Poles at* z = 1

(a) Position error $= \dfrac{1}{1 + G(z)}$

(b) Velocity error $= \infty$

(c) Acceleration error $= \infty$

Type 1: *One Pole at* z = 1

(a) Position error $= 0$

(b) Velocity error $= \dfrac{T(z-1)}{G(z)} = \dfrac{T}{K_v}$

(c) Acceleration error $= \infty$

Type 2: *Two Poles at* z = 1

(a) Position error $= 0$

(b) Velocity error $= 0$

(c) Acceleration error $= \dfrac{T^2(z-1)^2}{G(z)} = \dfrac{T^2}{K_a}$

17.3. Nonlinear Systems

17.3.1. Basic Definitions

GENERAL CHARACTERISTICS OF NONLINEAR SYSTEMS

(a) Superposition does not hold.

(b) Sinusoidal inputs do not necessarily produce sinusoidal outputs.

(*c*) System stability may depend on input amplitude or frequency.

(*d*) Instability may be exhibited by the presence of limit cycles, *i.e.,* oscillations of constant amplitude and arbitrary waveform.

(*e*) Subharmonics can be generated.

PHASE PLANE. A graphical determination of system performance in terms of state variables (output or error position, velocity, etc.).

SINGULAR POINT. A singular point is a point of unusual behavior in the phase plane.

STRUCTURAL STABILITY. If the system is described by

$$\dot{x} = P(x, y)$$

$$\dot{y} = Q(x, y)$$

small perturbations of the coefficients P and Q will not cause a change in the behavior of a structurally stable system. (Saddle points and centers are characteristic of structurally unstable systems.)

STABILITY (AROUND AN EQUILIBRIUM POINT)

(*a*) Stable in Liaponov sense: The state variables are bounded in a region for any initial disturbance in that region, and the upper bound is a function of the initial disturbance.

(*b*) Asymptotic stability: In addition to being stable in the Liaponov sense, the system approaches the equilibrium point asymptotically.

(*c*) Global stability: The system is asymptotically stable over the entire phase space.

LIMIT CYCLE. A limit cycle is an oscillation of constant amplitude and arbitrary waveform. There are two basic types of limit cycles.

(1) Stable limit cyçle: If a solution is started inside or outside the limit cycle, the system will approach the limit cycle. A stable limit cycle must enclose either an unstable limit cycle or an unstable equilibrium point.

(2) Unstable limit cycle: If a solution is started inside or outside the limit cycle, the system will diverge from the limit cycle. An unstable limit cycle must enclose a stable limit cycle or a stable equilibrium point.

HARD EXCITATION. Hard excitation is excitation such that the system is outside an unstable limit cycle.

CHARACTERISTIC EXPONENT OF LIMIT CYCLE. If n_0 is the perpendicular distance of a phase-plane path from the limit cycle, and n is the perpendicular distance of the path one cycle later, in the neighborhood of the limit cycle,

$$n = n_0 e^{a_0}$$

a_0 is the characteristic exponent; if $a_0 < 0$, then there is a stable limit cycle; and if $a_0 > 0$, then there is an unstable limit cycle.

FIRST CANONIC FORM OF DIFFERENTIAL EQUATIONS. Consider the system of Fig. 17-39. Let the following hold:

$$r = 0 \text{ (no input for } t > 0)$$

$$y = f(x) \text{ (description of nonlinear element)}$$

Then the system nonlinear element can be described by

$$\frac{dz_i}{dt} = \lambda_i z_i + f(x) \qquad i = 1, 2, \ldots, n$$

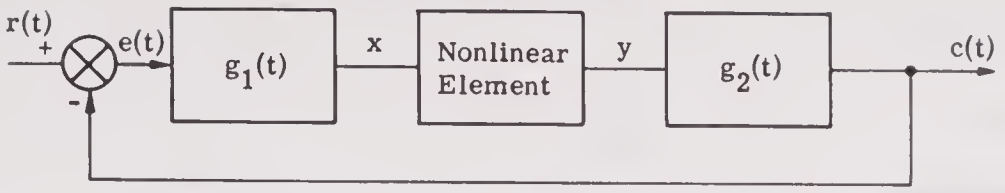


FIG. 17-39. Basic nonlinear feedback system.

$$x = \sum_{i=1}^{n} \alpha_i z_i$$

$$\frac{dx}{dt} = \sum_{i=1}^{n} \beta_i z_i - rf(x)$$

where λ_i are the poles of $G_1(s)G_2(s)$

α_i are the negatives of the residues of $G_1(s)G_2(s)$

$$\beta_i \; \alpha_i \lambda_i r = -\sum_{i=1}^{n} \alpha_i \qquad i = 1, 2, \ldots, n$$

Note that this description is valid only for systems in which the poles of $G_1(s)G_2(s)$ are simple and in which the number of zeros of $G_1(s)G_2(s)$ is less than the number of poles.

17.3.2. Methods of Analyzing Nonlinear Systems

17.3.2.1. *Time-Domain Analysis.* For systems whose nonlinearity can be simply described by low-order polynomials, (*e.g.,* straight-line approximations or simple curves), the system performance can be determined in a piecewise fashion. First write the system equations for operation in each mode of nonlinearity behavior. Then, for the given input, obtain a piecewise solution, using the system equation appropriate to the nonlinear behavior. To evaluate constants, use the output value at the end of each time interval as the initial condition for the interval which follows.

17.3.2.2. *Consideration of the Differential Equation*

BENDIXSON'S THEOREM. The system equation is written in terms of the state variables, i.e.,

$$\dot{x} = P(x, y)$$

$$\dot{y} = Q(x, y)$$

Then if

$$\frac{\partial P}{\partial_x} + \frac{\partial Q}{\partial_y} \neq 0$$

and is of constant sign in a region, there is no limit cycle in that region.

EXISTENCE OF LIMIT CYCLES FOR SECOND-ORDER SYSTEM. Consider the equation

$$\ddot{x} + f(x)\dot{x} + g(x) = 0$$

The following are sufficient but not necessary conditions for the existence of limit cycles.

(*a*) $f(x)$ and $g(x)$ are analytic (can be expressed as power series)

(*b*) $-g(x) = g(-x)$

(*c*) $xg(x) > 0$

(*d*) $f(x) = f(-x)$

(*e*) $f(0) < 0$

Define the functions

$$F(x) = \int_0^x f(x)\,dx$$

$$G(x) = \int_0^x g(x)\,dx$$

then:

(*f*) $F(x) \to \infty$ as $x \to \infty$

(*g*) $G(x) = 0$ has a unique root at $x = a$, where $a > 0$, and is monotone-increasing for $x > a$.

17.3.2.3. *Describing Function Analysis*

INTRODUCTION. In applying this method it is assumed that there is only one nonlinear element in the system, that the nonlinear element is time invariant, and, if the input is a sinusoid, that the only significant component of the output is the fundamental component; no subharmonics are generated.

Describing functions are defined by

$$N = \frac{\text{Amplitude of the fundamental at the output, } c}{\text{Amplitude of input sinusoid, } r} = \left|\frac{c_1}{r_1}\right|$$

The third harmonic of the output usually causes the largest error [2], hence $|c_3/c_1|$ measures the accuracy of the describing function in predicting the true output.

OPEN-LOOP DESCRIBING FUNCTION. The open-loop describing function N_0 can be found by first determining the nature of the nonlinearity and then hypothetically (or experimentally) inserting an input $c = X \sin \omega t$. Then N_0 is the plot of amplitude of fundamental of the output *vs.* X. (For examples of open-loop describing functions, see Fig. 17-40 through 17-43.)

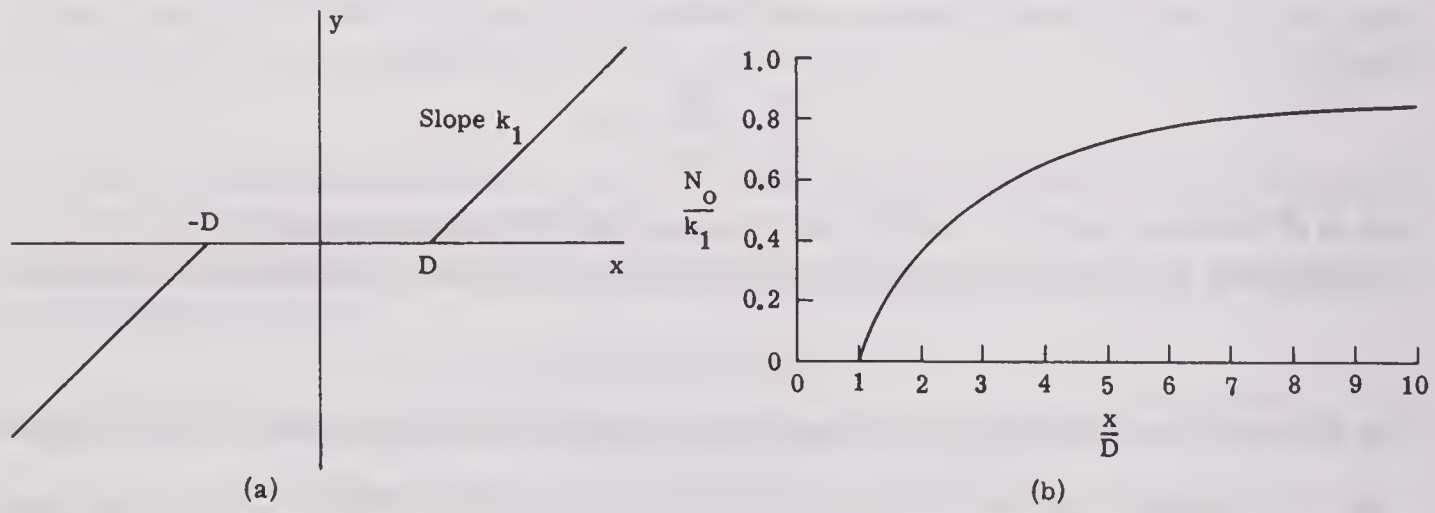


FIG. 17-40. Open-loop describing function for nonlinear element with dead zone. (*a*) Description of nonlinear element behavior. (*b*) Describing function.

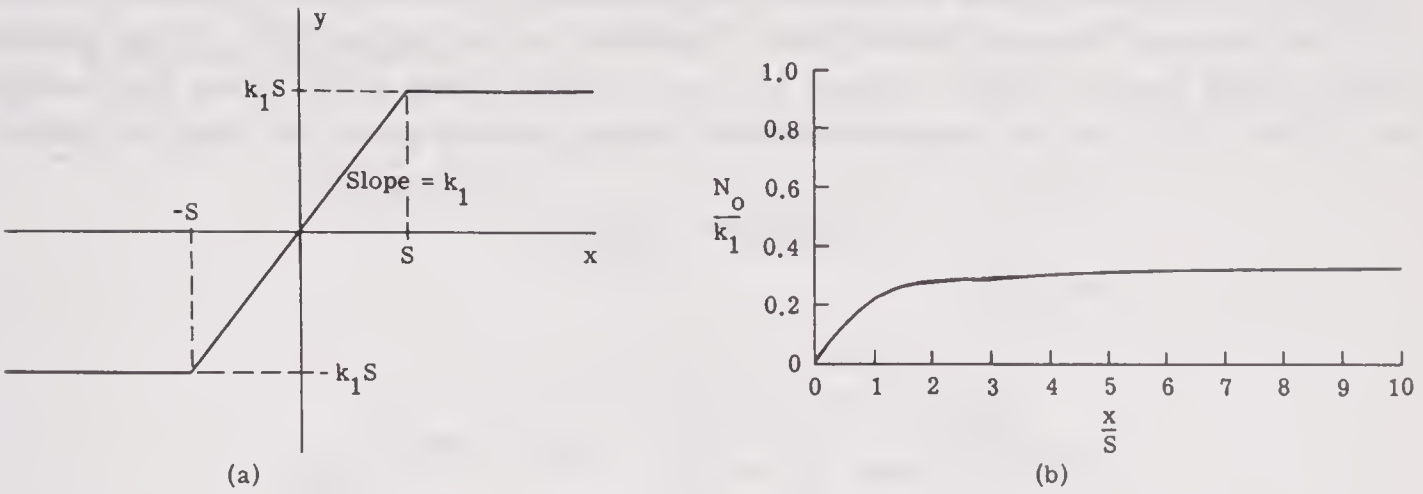


FIG. 17-41. Open-loop describing function for nonlinear element with saturation. (*a*) Description of nonlinear element behavior. (*b*) Describing function.

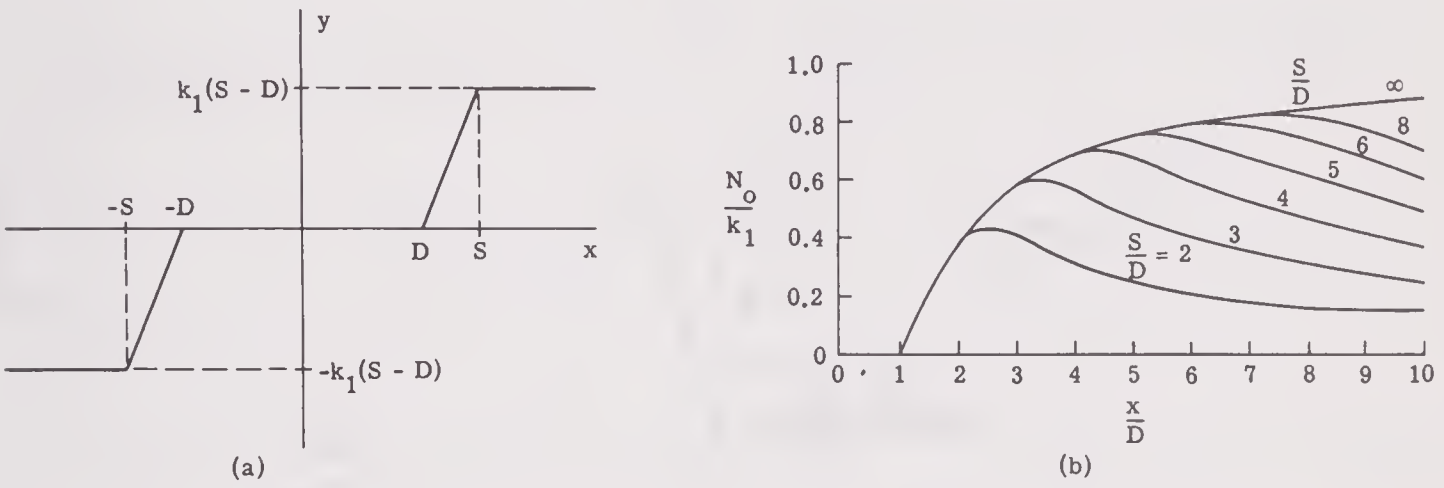


FIG. 17-42. Open-loop describing function for nonlinear element with dead zone and saturation. (*a*) Description of nonlinear element behavior. (*b*) Describing function.

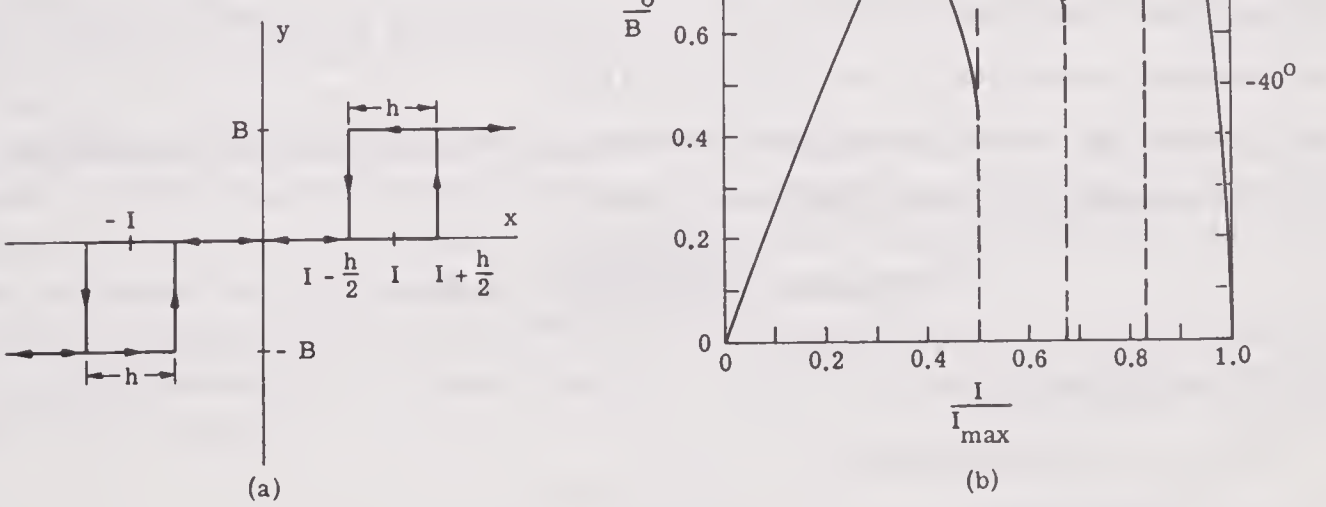


FIG. 17-43. Open-loop describing function for nonlinear element with hysteresis. (*a*) Description of nonlinear element behavior. (*b*) Describing function.

STABILITY ANALYSIS FROM DESCRIBING FUNCTIONS. If the system transfer function is $KG(s)$, make a Nyquist plot of $G(s)$. Consider the gain to be N_0K. Then plot the critical point locus, $-1/N_0K$, where N_0 may include phase shift. (See, for example Fig. 17-44.) If N_0 is frequency-dependent, several critical point loci must be plotted.

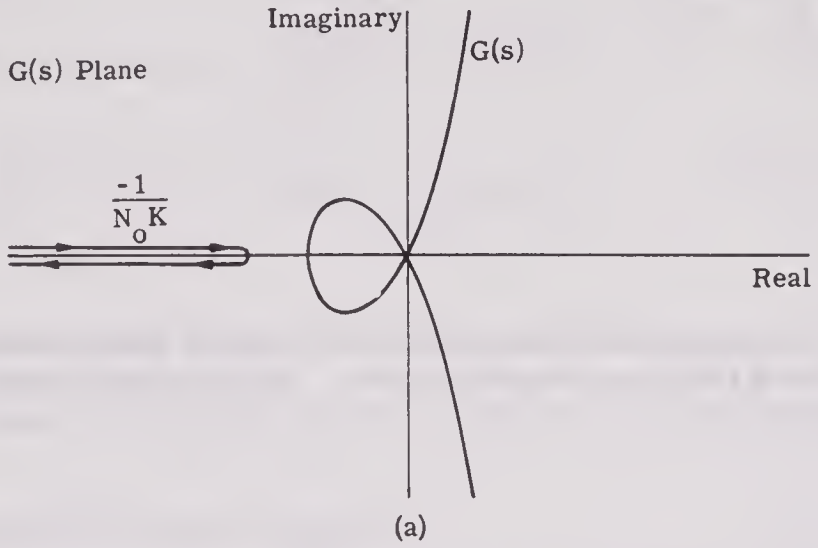


FIG. 17-44. Example of stability analysis using describing functions. (*a*) Stable for all input signals. (*b*) System with limit cycles: for input amplitudes less than *a*, the system is stable. For input amplitudes greater than *a*, the system has a stable limit cycle at *b*.

CLOSED-LOOP DESCRIBING FUNCTION N_c. N_c can be found as follows (see Fig. 17-45):

(*a*) Assume a frequency ω and calculate $KG(\omega)$

(*b*) Assume an error voltage and determine y from the describing function, N_0.

(*c*) The closed-loop system gain at ω is then

$$G_c(j\omega) = \frac{N_0 KG(j\omega)}{1 + N_0 KG(j\omega)} = Me^{jN}$$

(*d*) $C_0 = EN_0 KG(j\omega)$

$R_i = (1 + N_0 KG(j\omega)E\, e^{j\omega t}$

$$N = \left|\frac{N_0 KG(j\omega)}{N_0}\right| E$$

(*e*) Repeat for different E's and plot M and N *vs.* input amplitude A.

(*f*) Repeat for different ω.

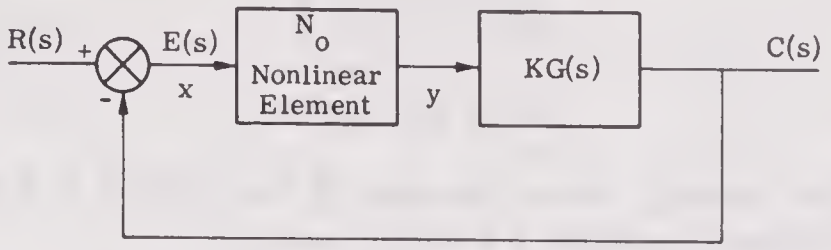


FIG. 17-45. Closed-loop system for describing function analysis.

EXAMPLE OF CLOSED-LOOP DESCRIBING FUNCTION. Figures 17-46 and 17-47 show the closed-loop describing function for a second-order system with torque saturation.

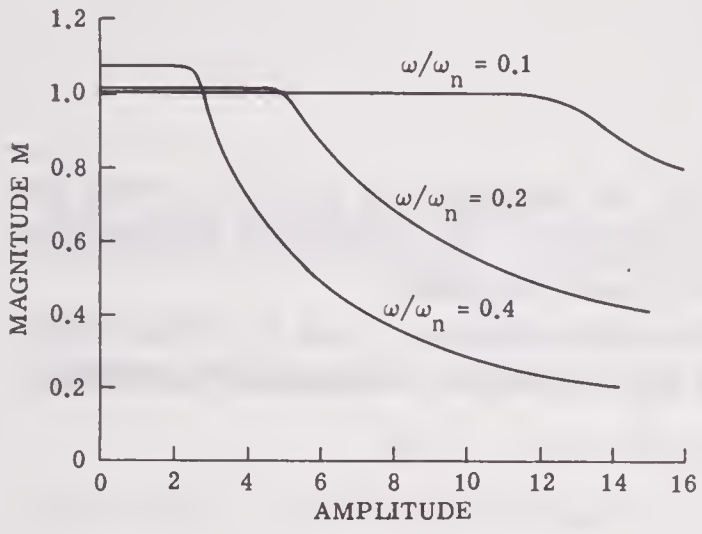


FIG. 17-46. Magnitude M of closed-loop describing function of second-order control system for nonlinear element with saturation.

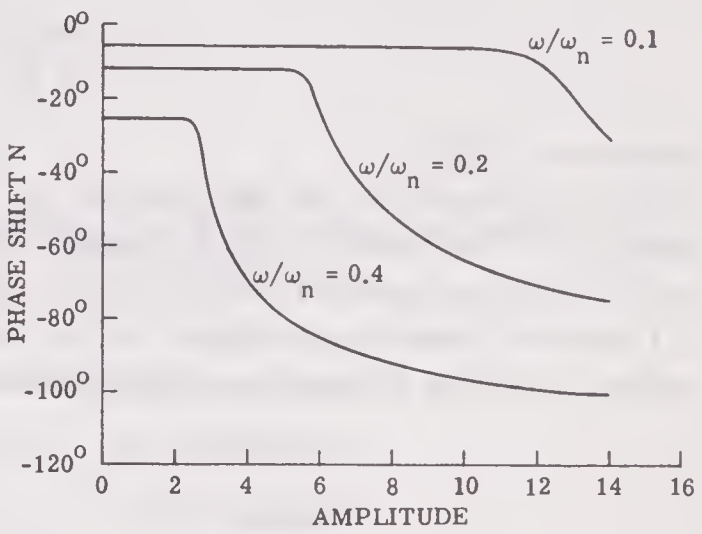


FIG. 17-47. Phase shift N of closed-loop describing function of second-order control system for nonlinear element with saturation.

17.3.2.4. *Phase-Plane Analysis*

INTRODUCTION.. Phase-plane analysis is most useful for second-order systems; it can be used to study transient behavior subject only to initial conditions (*i.e.*, no other excitation); only time-invariant systems can be considered. The performance is plotted with state variables as the coordinates. In general, this technique is not applicable to systems of higher order than 2 or 3, but by extending the concept it can provide insight into the performance of systems of any order.

OBTAINING PHASE PORTRAIT

Direct Solution of Differential Equations (for Linear Systems). The differential equation can be solved to obtain x, which is then differentiated to obtain $\dot{x}$. t is then eliminated between x and $\dot{x}$. Better results are usually obtained if the origin is moved to the maximum-velocity portion of the nonlinear characteristic.

Solution of Differential Equation for $\ddot{x}$. $y = \dot{x}$ can be substituted into the equation and the equation divided by $\dot{x}$. Noting that $\dot{y}/y \rightarrow dy/dx$, the new equation can often be solved. For example:

$$\ddot{x} + \omega_n^2 x = 0$$

$$\dot{y} + \omega_n^2 x = 0$$

$$\frac{\dot{y}}{y} + \frac{\omega_n{}^2 x}{y} = 0$$

$$\frac{dy}{dx} + \omega_n{}^2 \frac{x}{y} = 0$$

$$\frac{y^2}{\omega_n{}^2} + x^2 = K^2$$

The isoclines (lines through which the slope of the solution is a constant) are then determined. For example, consider $\ddot{x} + a_1{}^1 x + a_0 x = 0$

Let $\dot{x} = y$

$$\dot{y} = -a_1 y - a_0 x$$

$$\frac{\dot{y}}{\dot{x}} = \frac{dy}{dt} \cdot \frac{dt}{dx} = \frac{dy}{dx} = -\frac{a_1 y - a_0 x}{y}$$

or

$$y = -\frac{a_0}{s + b_1} x$$

where $s = dy/dx$.

This is the equation for the isocline on which the solution has slope s. Next, the direction of the solution can be determined, since $\dot{x} = y$. If y is positive, $\dot{x}$ is positive, and the solution moves toward the right in the upper-half plane.

Types of Phase-Plane Plots. If, for a second-order system, λ_1 and λ_2 are the eigenvalues, then the phase-plane plots shown in Fig. 17-48 are summarized as follows:

Unstable node	$\lambda_1 > 0, \quad \lambda_2 > 0, \quad \lambda_1 > \lambda_2$
Stable node	$\lambda_1 < 0, \quad \lambda_2 < 0$
Saddle point	$\lambda_1 \lambda_2 < 0$
Unstable focus	$\lambda = p + iq \qquad p > 0$
Stable focus	$\lambda = p + iq \qquad p < 0$
Center	$\lambda_1 = \lambda_2 = p + iq \qquad p = 0$

Time Response from Phase-Plane. The time response can be formed simply by forming

$$t_1 - t_2 = -\int_{x_1}^{x_2} \frac{1}{y}\, dx$$

where $y = dx/dt$ (the ordinate). Note that this equation can also be used in the error plane, where

$$t_1 - t_2 = -\int_{e_1}^{e_2} \frac{1}{\dot{e}}\, de$$

17.3.2.5. *Stability Analysis by Liaponov's Second Method*

Introduction. This method can be applied to higher-order systems if the differential equation is written in the first canonic form. However, this closed-loop system must have only one nonlinear element, and the nonlinear element must satisfy

$$\int_0^x f(\alpha)\, d\alpha \geq 0$$

for the closed-loop system with a single nonlinear element.

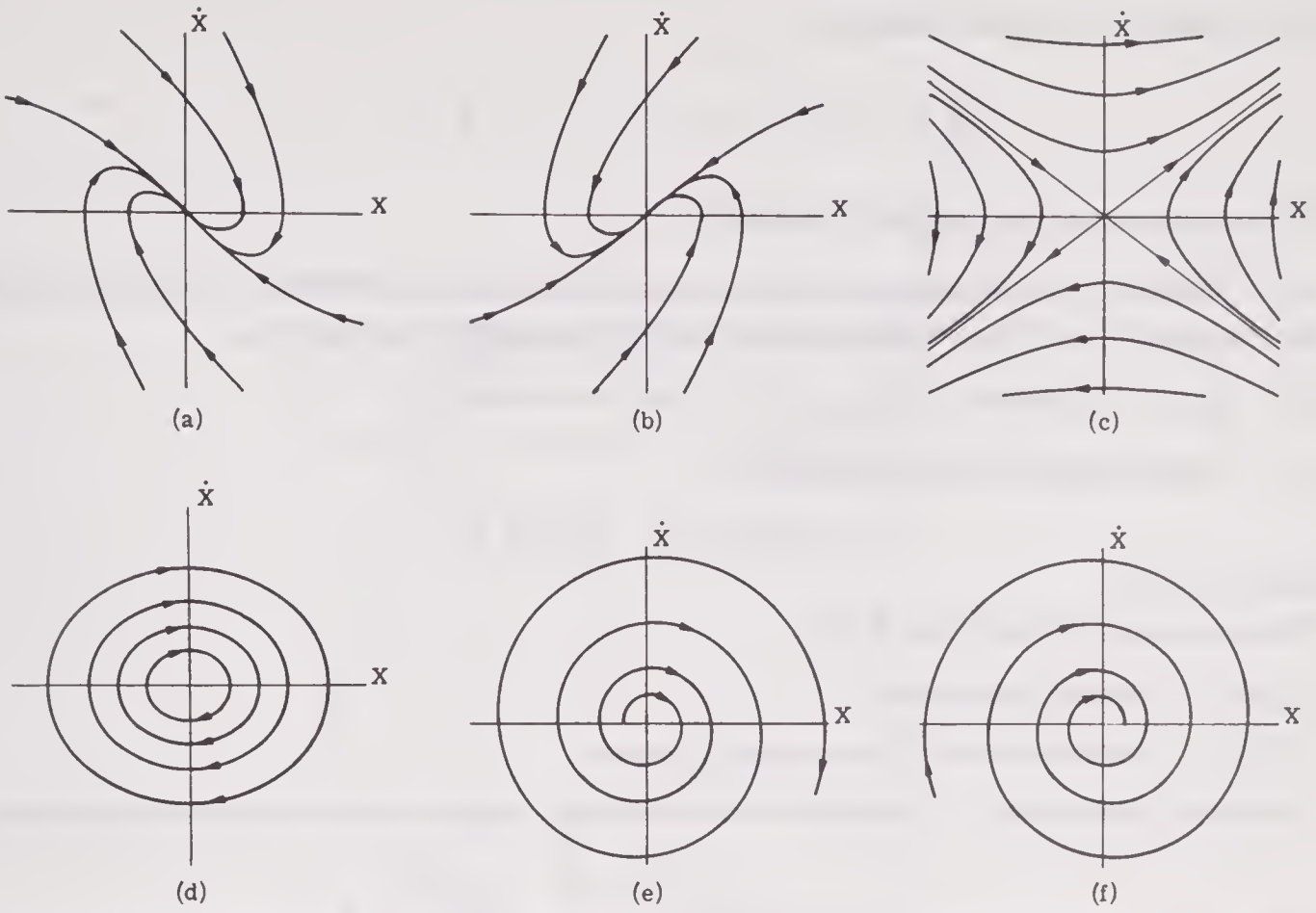


FIG. 17-48. Phase-plane plots for second-order system. (*a*) Stable node. (*b*) Unstable node. (*c*) Saddle point. (*d*) Center. (*e*) Unstable focus. (*f*) Stable focus.

LIAPONOV'S THEOREM. Consider the autonomous system

$$y_i = Y_i(y_1, y_2, \ldots, y_n)$$

If there exists a real-valued function $V(y_1, y_2, \ldots, y_n)$ such that.

(*a*) $V(y_1, y_2, \ldots, y_n)$ has continuous first partial derivatives

(*b*) V is positive definite, *i.e.*, $V > 0$ for any $|y_i| > 0$, $V(0) = 0$

(*c*) $\lim_{|y_i| \to \infty} V(y_1, y_2, \ldots, y_n) = \infty$ for all $|y_i| \to \infty$

then the equilibrium state of the system is as follows:

(*a*) Stable, if there is some region $V < k$, (where k is some positive constant) such that

$$\frac{dV}{dt} = \sum_{i=1}^{n} \frac{\partial V}{\partial y_i} \frac{dy_i}{dt} \leqslant 0$$

i.e., dV/dt is negative-semidefinite.

(*b*) Asymptotically stable, if $dV/dt < 0$ in the region $V < k$, or if $dV/dt \leqslant 0$ and $dV/dt = 0$ is not a nontrivial solution of

$$\vec{y}_i = \vec{Y}_i\,(y_1, y_2, \ldots, y_n)$$

(*c*) Globally stable if condition (*b*) holds for the entire phase space.

LIAPONOV FUNCTION FOR SYSTEMS DESCRIBED BY FIRST CANONIC FORM:

$$V = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} Y_i Y_k a_{ik} = a_{ki}$$

The stability equation becomes

$$2a_i \sum_{j=1}^{n} a_j/(\lambda_i + \lambda_j) = \alpha_i \qquad i = 1, 2, \ldots, n$$

The system is asymptotically stable if

(*a*) There is at least one solution such that the a_i are real for real λ_i's and are in complex conjugate pairs for corresponding complex conjugate pairs of λ_i.

(*b*) $Re\lambda_i < 0$ for all $i = 1, 2, \ldots, n$

(*c*) The nonlinear element satisfies

$$xf(x) \geq 0, \qquad f(0) = 0$$

for all $|x| > 0$

More examples are found in [7].

17.3.3. Specific Solutions

17.3.3.1. *Second-Order Conservative Systems*

GENERAL COMMENTS. The equation describing a second-order conservative system is

$$\ddot{x} + f(x) = 0$$

Second-order conservative systems exhibit no damping and are described in the phase-plane by

$$\dot{x} = y$$

$$\dot{y} = -f(x)$$

or

$$\frac{dy}{dx} = -\frac{f(x)}{y}$$

POTENTIAL FUNCTION. A function $v(x)$ can be defined as

$$v(x) = \int_0^x f(x)\, dx$$

Hence the system equation becomes

$$y^2/2 \quad + v(x) \quad = h$$

or

kinetic energy	+ potential energy	= total energy

BEHAVIOR NEAR AN EQUILIBRIUM POINT. If the second derivative of $v(x)$ is taken, then, if

$$\frac{d^2v(x_0)}{dx^2} = \frac{df(x_0)}{dx} > 0$$

The system exhibits center behavior, if

$$\frac{d^2v(x_0)}{dx^2} = \frac{df(x_0)}{dx} < 0$$

The system exhibits saddle behavior; and if (see Fig. 17-49)

$$\frac{d^2v(x_0)}{dx^2} = \frac{df(x_0)}{dx} = 0$$

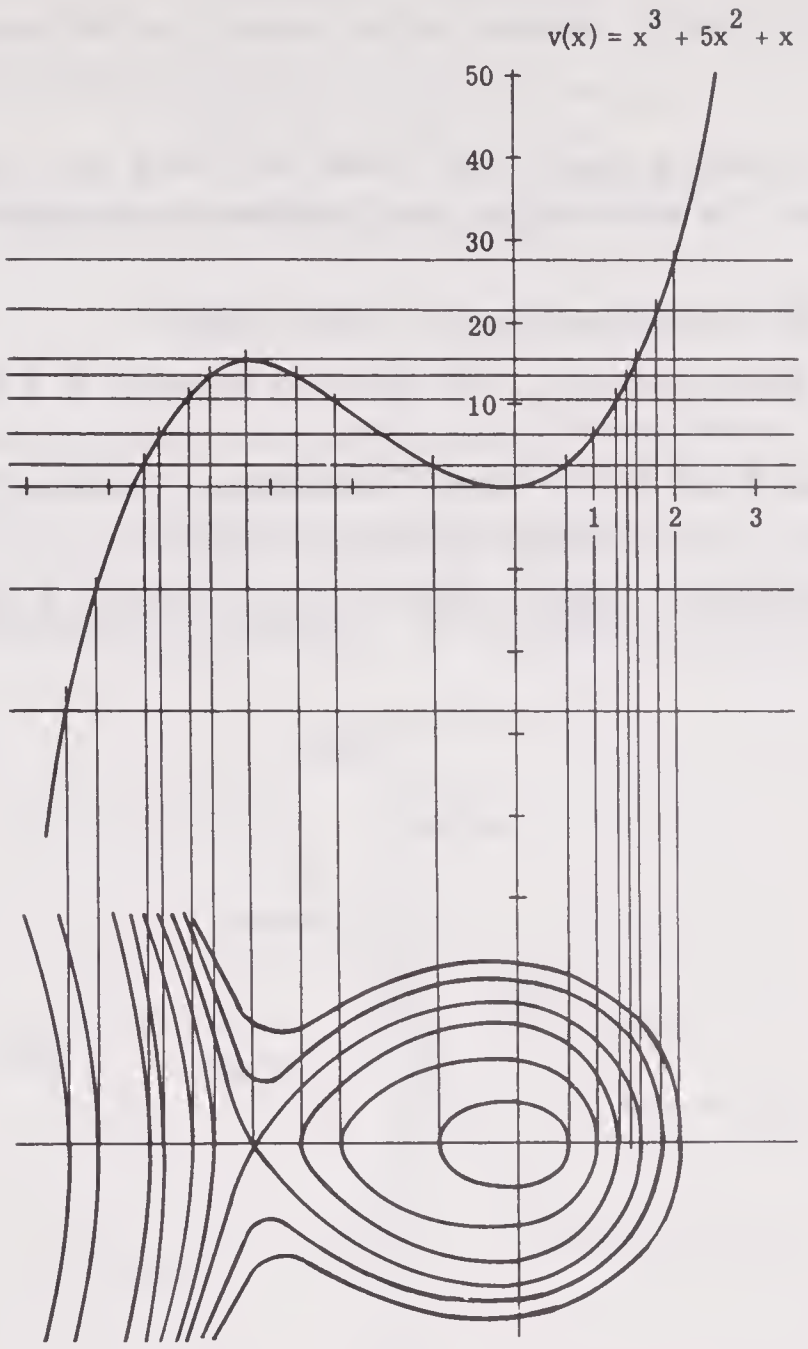


FIG. 17-49. Construction of phase-plane plot of $\ddot{x} + 3x^2 + 10x + 1 = 0$.

SKETCH OF PHASE PORTRAIT [2] FOR CONSERVATIVE SECOND-ORDER SYSTEM. In sketching the phase portrait, the singular points are located (maximums and minimums of v) and the nature of each point determined from the preceding discussion. Then the lowest energy path is constructed, then the next higher, etc. (see Fig. 17-49).

EFFECT OF VARYING A PARAMETER σ

(a) Consider the equation

$$\ddot{x} + f(x, \sigma) = 0$$

(b) Plot $f(x, \sigma) = 0$, which is a plot of equilibrium points.

(c) Shade the region where $f(x, \sigma) > 0$; then, for a given σ, the equilibrium points are given by the curve. The number and type of equilibrium points can be determined as follows: if, as x increases, $f(x, \sigma)$ changes from shaded to unshaded; the system exhibits saddle behavior; if the change is from unshaded to shaded, the system exhibits center behavior.

17.3.3.2. *Examples of Systems with Saturation*

INTRODUCTION. An error plane can be constructed as follows:

(a) Write the system characteristic equation in terms of

$$\epsilon = \theta_i - \theta_o$$

The linearity of this system depends on the linearity of the torque output T since

$$\ddot{e}(t) = T[\dot{e}(t), e(t), \int e(t)\,dt, \ldots]$$

(*b*) Determine switching lines which divide the plane into regions of linear and nonlinear operation (it is assumed that nonlinearities can be approximated by straight lines).

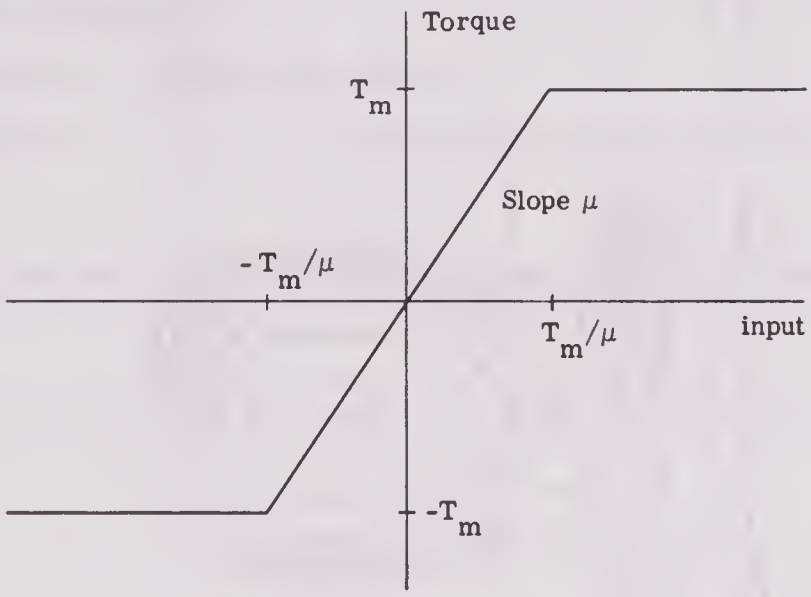
(*c*) Determine the qualitative behavior in each region.

(*d*) Sketch the phase portrait. Note that the behavior of a saturated system is parabolic.

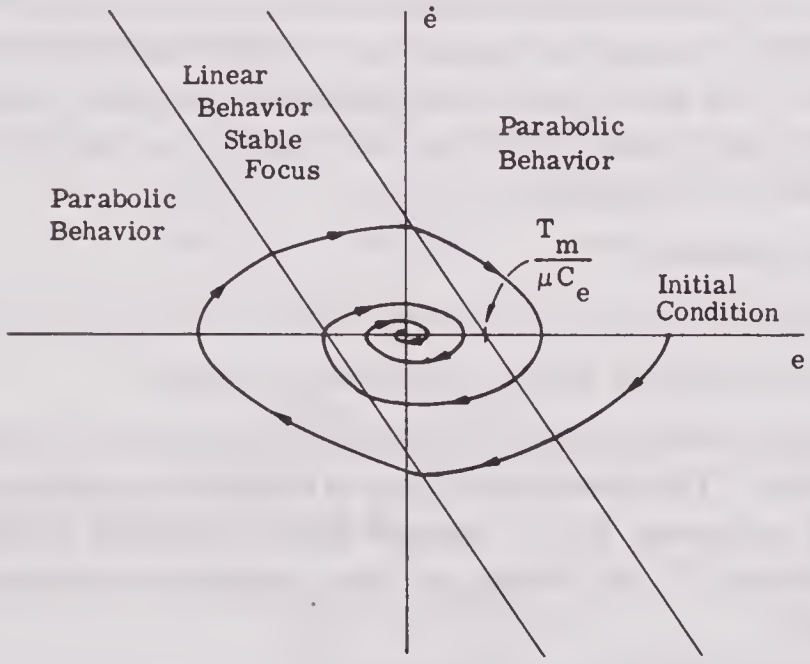
Error Plus Error-Rate System with No Damping. Consider the system

$$I\ddot{e}(t) = \mu\,[e(t) + C_e\dot{e}(t)]$$

where C_e is the error-rate coefficient and the torque saturation curve is as shown in Fig. 17-50(*a*).



(a)



(b)

Fig. 17-50. Phase-plane for system with torque saturation. (*a*) Description of nonlinear element. (*b*) Phase portrait.

If there is no acceleration input, $I\ddot{r}(t) = 0$ and

$$I\ddot{e}(t) = -\mu[C_e\dot{e}(t) + e(t)]$$

from which

$$T = \mu[C_e\dot{e}(t) + e(t)]$$

and, integrating

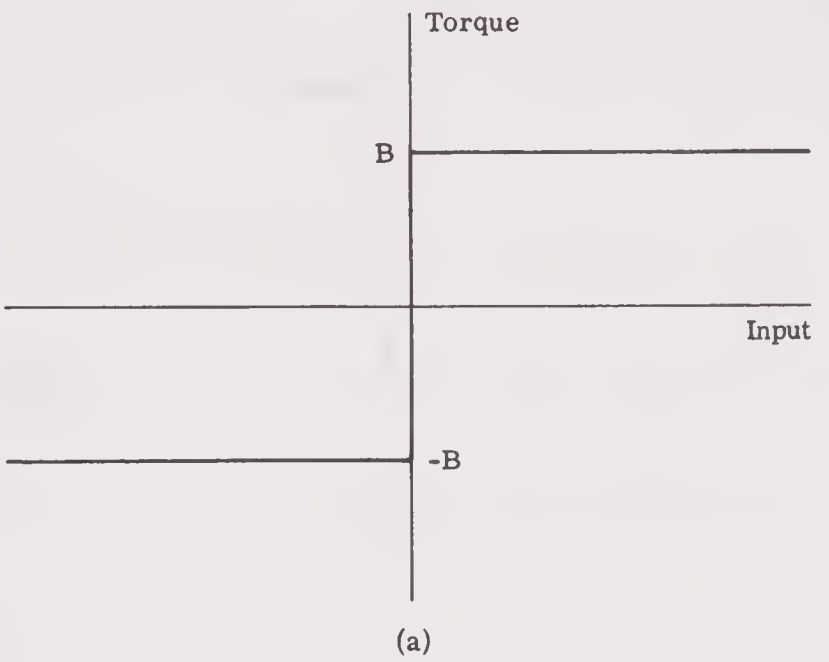
$$\dot{e}(t) = -\frac{1}{C_e}\,e(t) + \frac{T}{\mu C_e}$$

Therefore, the switching lines are straight, representing torques of $\pm T_m$ with slope $-1/C_e$ and abscissa intercept $\pm T_m/\mu C_e$ (see Fig. 17-50*b*).

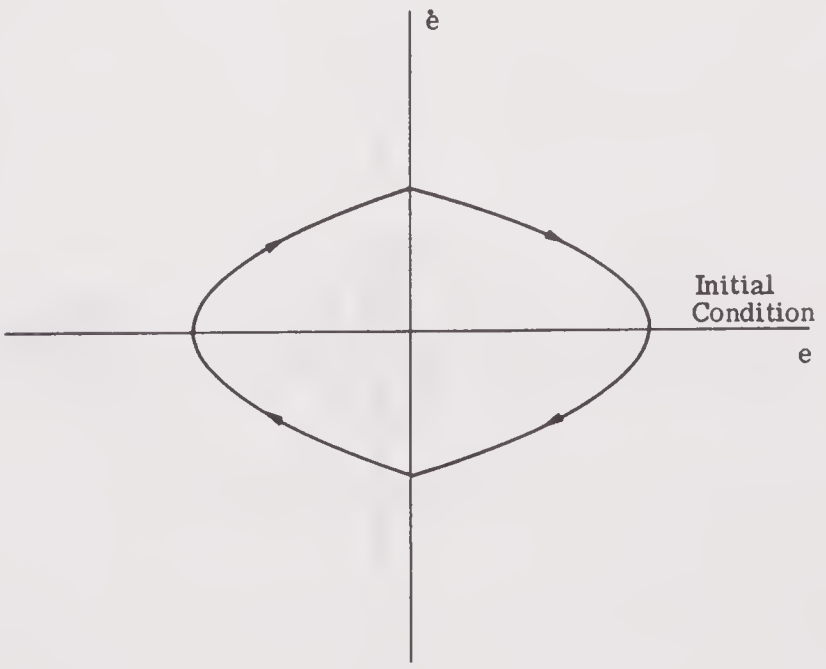
The behavior (see Fig. 17-50*b*) is linear between the saturated torque lines and parabolic on either side of this region. The phase portrait is sketched in Fig. 17-50*b*.

BANG-BANG SYSTEM WITH NO DAMPING. In this system the torque is described (see Fig. 17-51*a*) by

$$T = B\,e(t)/|e(t)| = \pm B$$



(a)



(b)

FIG. 17-51. Phase-plane for bang-bang system without damping. (*a*) Description of nonlinear element behavior. (*b*) Phase portrait.

and the system equation is

$$I\ddot{c}(t) = B\ e(t)/|e(t)|$$

There is no region of linear operation; the switching line is the ordinate, and the behavior on both sides of the switching lines is parabolic (see Fig. 17-51*b*).

BANG-BANG SYSTEM WITH COULOMB DAMPING. The torque is described by

$$T = B\ e(t)/|e(t)| - C\ \dot{e}(t)/|\dot{e}(t)|$$

(see Fig. 17-52*a*) so the system behavior is

$$I\ddot{c}(t) = B\ e(t)/|\dot{e}(t)| + C\ \dot{e}(t)/|\dot{e}(t)|$$

Again there is no linear operation, and the behavior is parabolic in each quadrant:

(*a*) Quadrant 1: $T = B + C$

(*b*) Quadrant 2: $T = -B + C$

(*c*) Quadrant 3: $T = -B - C$

(*d*) Quadrant 4: $T = B - C$

(See Fig. 17-52*b*).

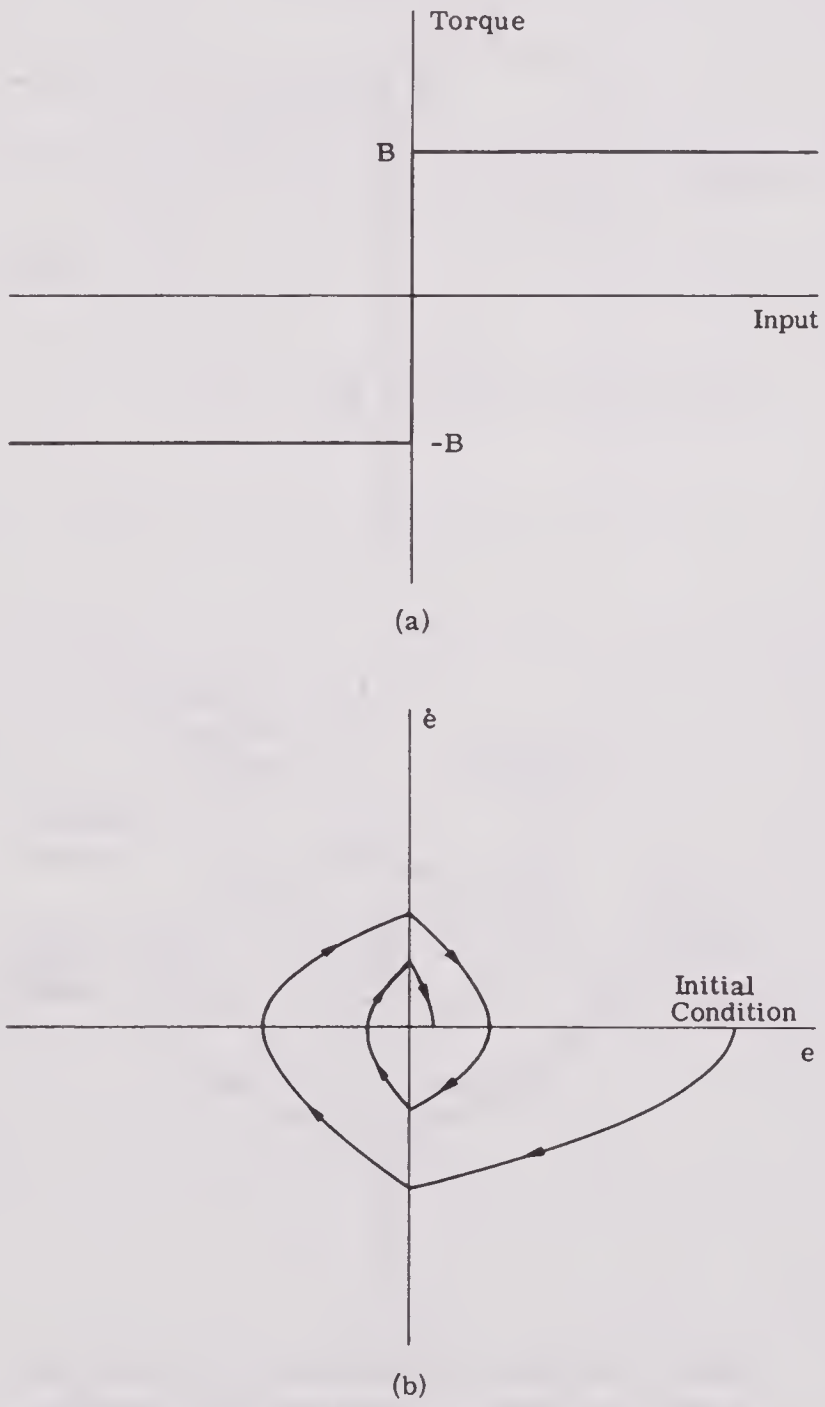


FIG. 17-52. Phase-plane for bang-bang system with coulomb damping. (*a*) Description of nonlinear element behavior. (*b*) Phase portrait.

BANG-BANG SYSTEM WITH VISCOUS DAMPING. The torque is described as

$$T = B\ e(t)/|e(t)| + f\dot{e}(t)$$

(see Fig. 17-53); the system equation is

$$I\ddot{c}(t) = B\ e(t)/|e(t)| + f\dot{e}(t)$$

There is no region of linear operation, and the behavior is parabolic:

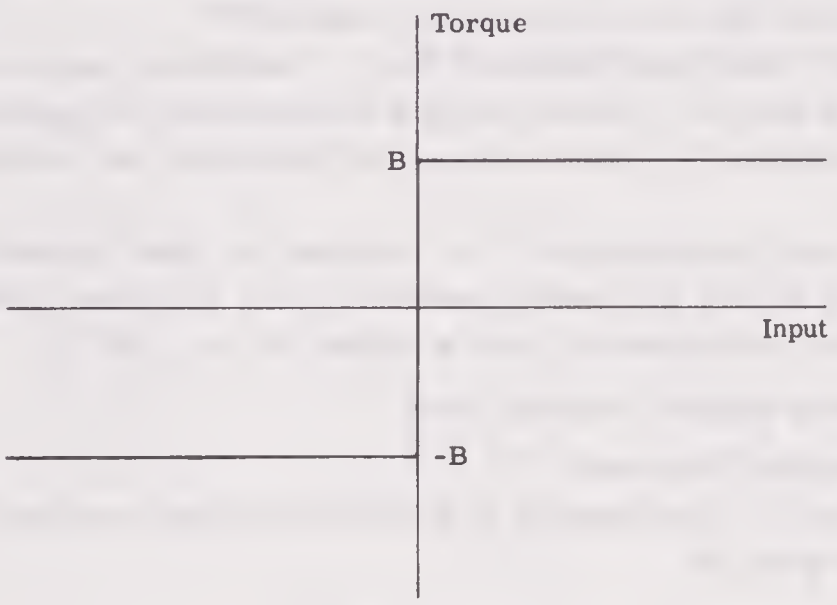
(*a*) Quadrant 1: $T = B + f\dot{e}(t)$

(*b*) Quadrant 2: $T = -B + f\dot{e}(t)$

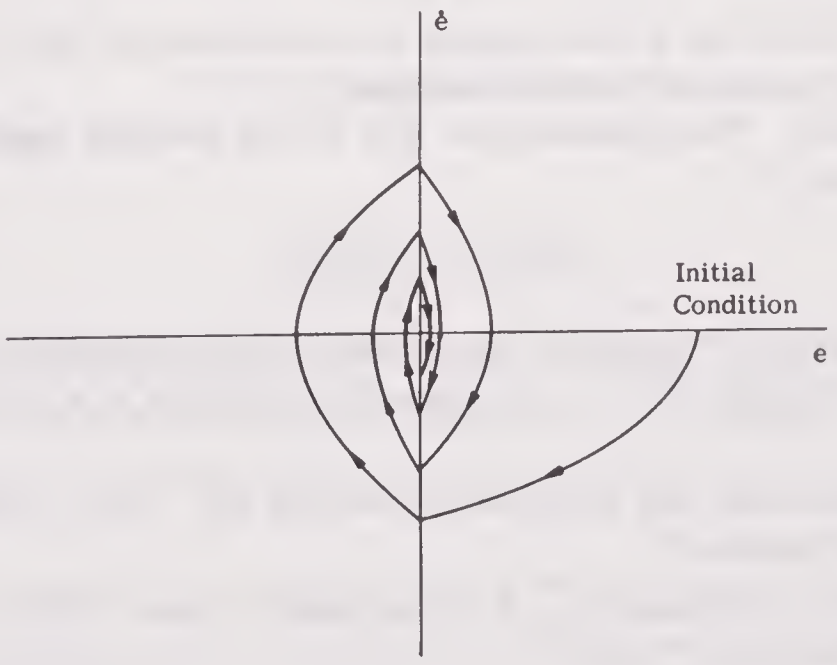
(*c*) Quadrant 3: $T = -B + f\dot{e}(t)$

(*d*) Quadrant 4: $T = B + f\dot{e}(t)$

That is, the system is highly damped for large error rates, and lightly damped for small error rates (see Fig. 17-53).



(a)



(b)

FIG. 17-53. Phase-plane for bang-bang system with viscous damping. (*a*) Description of nonlinear element behavior. (*b*) Phase portrait.

17.4. Design Methods

17.4.1. Gain Adjustment

GAIN ADJUSTMENT FROM THE POLAR PLOT. System performance can be strongly influenced by gain adjustment; the proper adjustment for desired operation can be derived from using a Nyquist plot as follows: Draw a Nyquist plot of $G(j\omega)$ [or of $G^*(j\omega)$ for a sampled-data system]. Draw a ψ line through the origin at $\psi = \sin^{-1}(1/M_p)$, where M_p is the desired peaking of the transfer function. Draw a circle centered on the negative real axis, tangent to both the ψ line and the $G(j\omega)$ locus. Determine K from

$$K = \frac{M_p^2}{c(M_p^2 - 1)} = \frac{1}{d}$$

where c = the distance from the origin to the center of the circle

d = the distance from the origin to the negative real-axis intercept of a perpendicular which passes through the point of circle and ψ-line tangency.

GAIN ADJUSTMENT FROM BODE PLOT. Proper gain adjustment is that which yields the desired gain and/or phase margin (usually about 30°).

GAIN ADJUSTMENT FROM LOG MODULUS PLOT. Variations in gain are reflected in a vertical shifting of the plot. Hence, using a Nichols chart, the gain which will result in tangency to the desired M circle is chosen by adjusting the vertical position of the plot.

COMMENTS ON GAIN ADJUSTMENT. An increase in gain decreases stability, decreases steady-state errors, increases transient errors, and lowers response time. In general, the result of an increase in gain is higher M_p, ω_n, and ω_p.

17.4.2. Cascade or Series Compensation

SIMPLE COMPENSATION NETWORKS

Phase Lead Network. This network (Fig. 17-54) provides positive phase shift. The transfer function is given by

$$K_c G_c(s) = \alpha \frac{s\tau_2 + 1}{s\alpha\tau_2 + 1}$$

where $\alpha = R_1/(R_1 + R_2)$

$$\tau_2 = R_2 C_2$$

The Nyquist and Bode plots of this network are also shown in Fig. 17-54. The phase shift obtained as a function of α is also indicated.

Phase Lag Network. This network (see Fig. 17-55) provides negative phase shift. The transfer function is

$$K_c G_c(s) = \frac{\tau_2 s + 1}{\tau_2 a s + 1}$$

Figure 17-55 illustrates Nyquist and Bode plots as well as the obtainable phase shift.

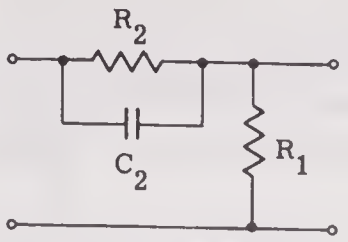
DESIGN USING NYQUIST PLOT. A compensating network can be designed by the following.

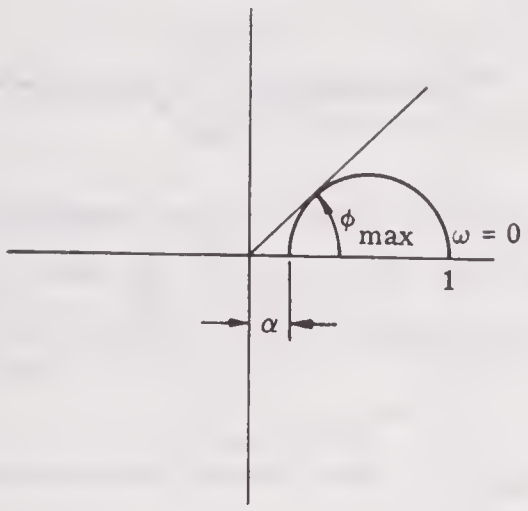
(*a*) Draw Nyquist plot and M circle for desired M_p. (M_p is approximately the transient response overshoot.)

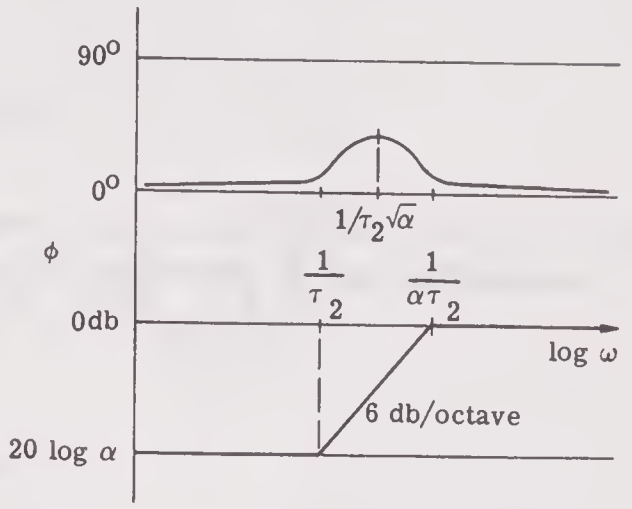
(*b*) Select resonant frequency ω_p of compensated system. This is related to the speed of response (see Sec. 17.1.3).

(*c*) If ω_p is the frequency of peak of the uncompensated systems, then the network must yield a phase shift of

$$\phi = \angle O\omega_p' - \angle O\omega_p$$







$$\phi_{max} = \tan^{-1}\frac{1}{\sqrt{\alpha}} - \tan^{-1}\sqrt{\alpha}$$

α	ϕ_{max}
0.1	53.2
0.2	41.4
0.3	32.0
0.4	25.0
0.5	19.4

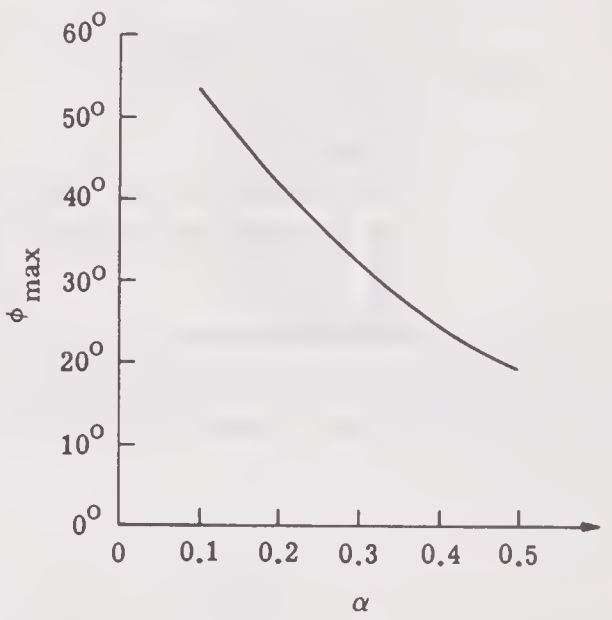


FIG. 17-54. Phase-lead network.

and a gain of

$$\mu = \overline{O\omega_p'} \Big/ \overline{O\omega_p}$$

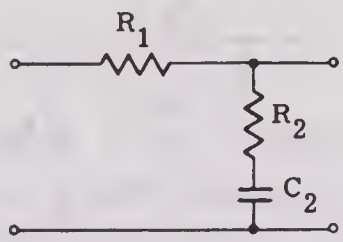
where $\overline{O\omega_p'}$ is the distance from the origin to ω_p' and $\overline{O\omega_p}$ is the distance from the origin to ω_p. Therefore, for a phase-lead network, Fig. 17-54 is used to determine α and for a phase lag network Fig. 17-55 is used to determine a; τ_2 is determined from (*b*) above.

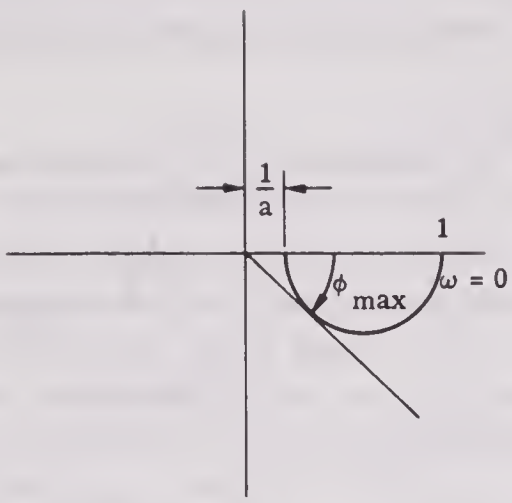
(*d*) The locus of the compensating network is plotted, using (τ_2, α) (or a) and μ, as derived.

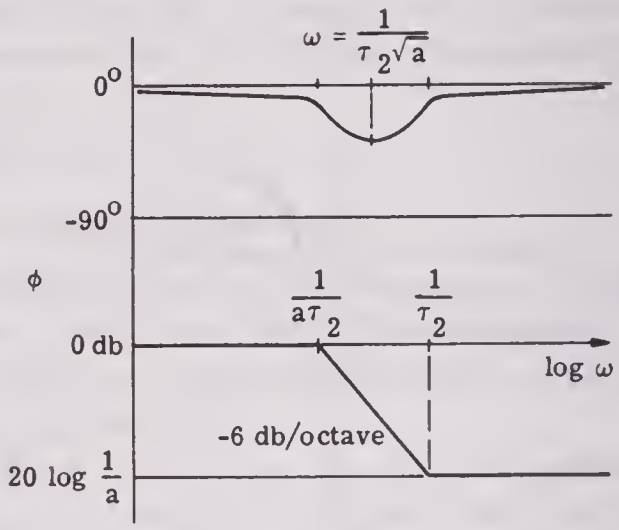
(*e*) By vector multiplication, the new system locus is constructed.

(*f*) Repeat the steps above as required for different ω_p and α (or a) until a satisfactory system is obtained.

In sampled-data systems it may be desirable to use a pulsed-data network for compensation. In this case, the procedure is as outlined above, but the z-plane is used for design.







$$\phi_{max} = \tan^{-1}\frac{1}{\sqrt{a}} - \tan^{-1}\sqrt{a}$$

a	ϕ_{max}
2	-9.73
3	-24.68
4	-33.43
5	-41.87
6	-45.58
10	-54.9

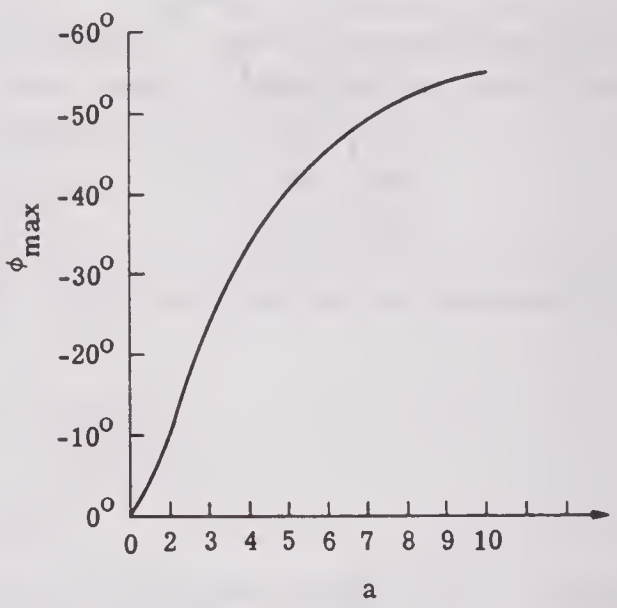


FIG. 17-55. Phase-lag network.

DESIGN USING LOG MAGNITUDE AND PHASE PLOTS. The method outlined below may be applied to both the log modulus and the Bode plots.

(*a*) Draw the Bode or log modulus plot of the uncompensated system and decide what type of compensation network is required.

(*b*) Determine the range of frequencies over which compensation is necessary.

(*c*) Draw the plot for the compensating network.

(*d*) Add the compensation plot to the system plot.

(*e*) Adjust the system gain as necessary to achieve desired performance.

(*f*) Repeat as necessary to obtain desired result.

OTHER COMPENSATION SCHEMES. Some additional compensation schemes follow.

First-Derivative Error Control. Although there is no change in steady-state error K_p, the transient response is faster. Application of first-derivative error control does not change the order of system and increases K_v without affecting the stability or phase margin excessively. It does, however, affect the apparent viscous damping (see Fig. 17-56).

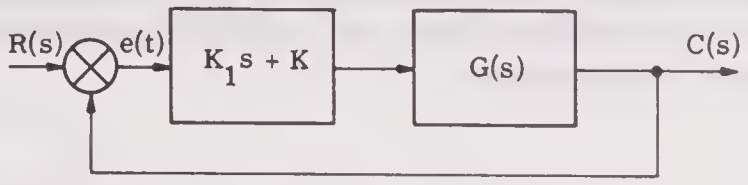


FIG. 17-56. First-derivative error control system.

First-Derivative Input Control. This method of compensation improves transient response, raises K_p, and K_v, and reduces the amplitude of the oscillations. It also affects the apparent viscous damping (see Fig. 17-57).

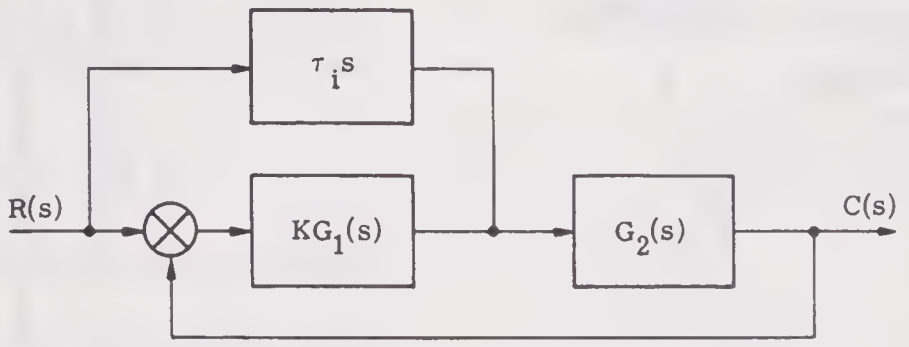


FIG. 17-57. First-derivative input control system. $G_1(s)$ is the controller. $G_2(s)$ is the controlled system.

First-Derivative Output Control. This system can reduce steady-state velocity error (even make it negative). Since the nature of transient response is changed, caution should be used in applying it. Again, the apparent viscous damping is affected (see Fig. 17-58).

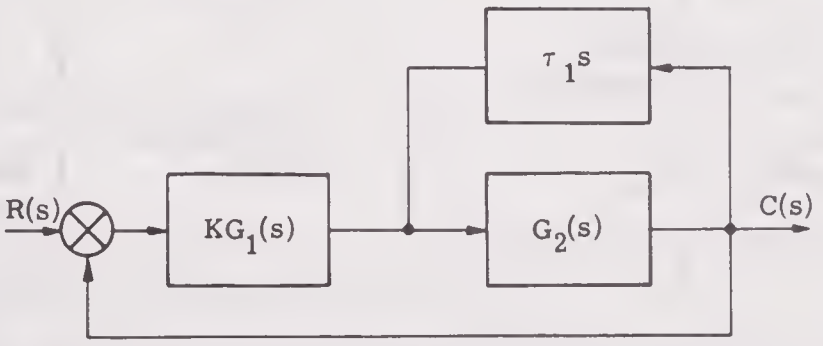


FIG. 17-58. First-derivative output control system.

Integral Control. This type of compensation reduces the steady-state error, which eventually becomes zero. Speed of response is increased but relative stability is decreased (see Fig. 17-59).

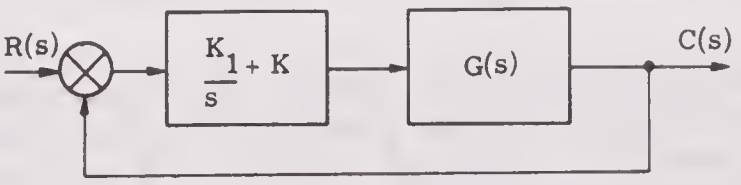


FIG. 17-59. Integral control system.

17.4.3. Root-Locus Method

17.4.3.1. *Introduction.* The effect of adding poles and zeros to a system is indicated in Fig. 17-60. Real-axis zeros tend to spread the loci faster and stabilize the system. Real-axis poles, on the other hand, tend to make the loci spread more slowly and curve toward instability.

The root-locus design method (after Guillemin in [2]) involves three steps: the closed-loop transfer function is determined, then the open-loop transfer function, and finally a compensation network is synthesized.

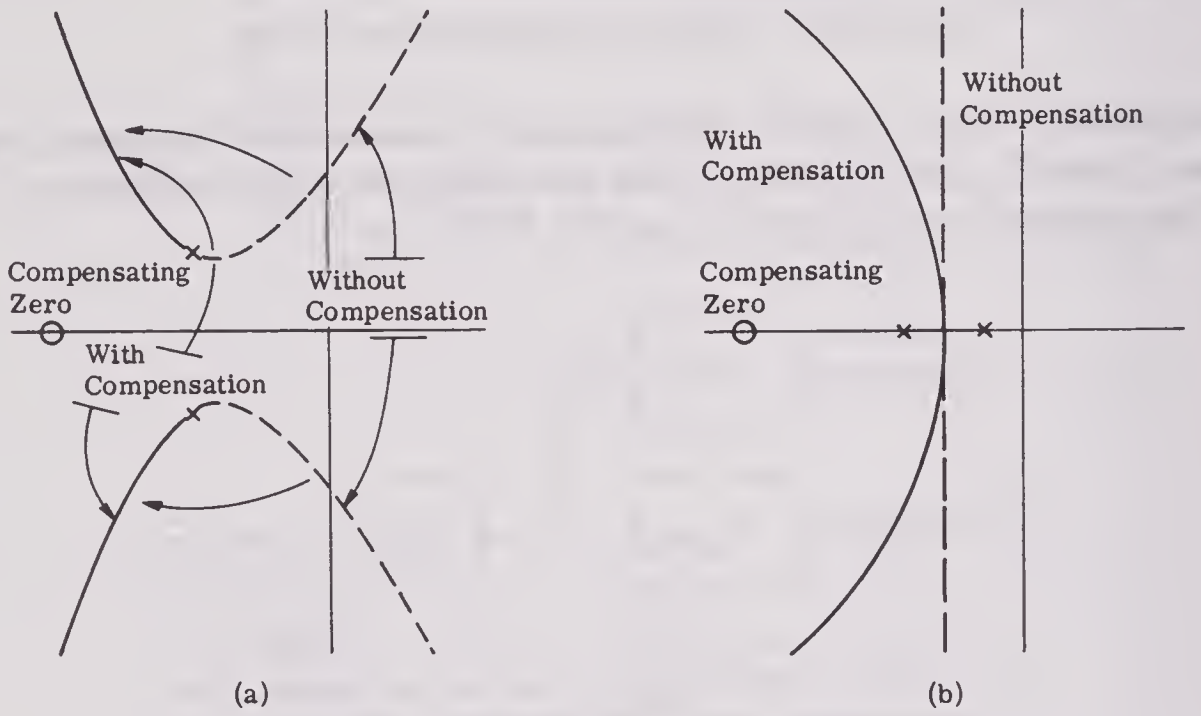


FIG. 17-60. Effect of additional poles and zeros. (*a*) Effect of zero on locus with complex poles. (*b*) Effect of zero on locus with real negative poles. (*c*) Effect of pole on locus with complex poles. (*d*) Effect of pole on locus with real negative poles.

17.4.3.2. *Determination of Closed-Loop Transfer Function.* Since the fixed elements of a system cannot be changed to bring about control, their characteristics must be determined by experimental or analytical techniques; these characteristics are then used to design an effective control system.

There are several basic relationships between the system specifications and the poles and zeros. If the system transfer function is written

$$\frac{C(s)}{R(s)} = K\frac{(s+z_1)(s+z_2)(s+z_m)}{(s+p_1)(s+p_2)(s+p_n)}$$

Then the basic relationships between the system specifications and the poles and zeros are

(1) $(p-z)_t \geq (p-z)_c$

where p is the number of poles

z is the number of zeros

(2) $$K_p = \frac{K\prod_{j=1}^{m} Z_j}{\prod_{j=1}^{n} p_j - K\prod_{j=1}^{m} Z_j}$$

where K_p is the position-error constant;

(3) $$\frac{1}{K_v} = \sum_{j=1}^{n}\frac{1}{p_j} - \sum_{j=1}^{m}\frac{1}{Z_i}$$

where K_v is the velocity-error constant;

(4) $$-\frac{2}{K_a} = \frac{1}{K_v 2} + \sum_{j=1}^{n}\frac{1}{p_j 2} - \sum_{j=1}^{m}\frac{1}{{Z_j}^2}$$

where K_a is the acceleration-error constant.

Example:

$$\frac{C(s)}{R(s)} = \frac{{\omega_n}^2}{s^2 + 2\omega_n s + {\omega_n}^2}$$

$$K_p = 1$$

$$K_v = \frac{2\xi}{\omega_n}$$

$$\frac{1}{K_a} = \frac{(1-4\xi^2)}{{\omega_n}^2}$$

(5) Delay time (for step input) $T_d \approx 1/K_v$ (exact if $g(t)$ is symmetrical around its centroid, *i.e.,* the system exhibits small overshoot).

(6) Rise time: (exact if $g(t)$ is symmetrical around its centroid, *i.e.,* the system exhibits small overshoot).

$$\frac{{T_r}^2}{2} = \frac{2}{K_a} - \frac{1}{{K_v}^2}$$

Method for Compensation Using Poles and Zeros. The object of this compensation method is to base control primarily on two control poles, *i.e.,* complex conjugate poles close to the imaginary axis. These poles and any significant zeros determine the system characteristics. In placing the control poles, other poles are placed either so far out from the imaginary axis, or so close, that they are not significant. In more complex systems, inner loops may have to be designed with other poles and zeros (see Sec. 17.4.5).

The procedure described above ensures that the open-loop poles lie on the negative real axis; hence, the compensation can be accomplished with passive RC networks.

POTENTIAL ANALOGY FOR K_v ADJUSTMENT. The object of this adjustment is to increase K_v so as to reduce errors in response to velocity inputs. This adjustment should be made without materially affecting the relative stability of the system.

If $K_p = -\frac{1}{2}\left(i.e., \frac{C(0)}{R(0)} = 1\right)$, the velocity constant K_v is analogous to the inverse of the electric field at the origin. Hence, poles are considered to be unit line charges (perpendicular to s-plane) and zeros are unit line charges of opposite polarity. Therefore, the field at any radius r from a pole or zero is $\pm 1/r$.

Using this analogy, zeros can be placed on the negative real axis in such a position as to reduce the field at the origin to 0; hence, an infinite K_v is obtained. This is lead control.

Note that a pole and zero could be placed close together so their net effect is negligible on the other poles and zeros, but close to the origin so that the effect on the field at the origin is great. If the pole is outside the zero, K_v is increased, while if the pole is inside the zero (closer to the origin), K_v is decreased. The increase in overshoot produced by this compensation is P_i/Z_i, and, because of the term $\epsilon^{-P_i t}$, a long-duration "tail" will be found in the response.

17.4.3.3. *Determination of Open-Loop Transfer Function.* Once $C(s)/R(s)$ has been obtained, the open-loop transfer function $G(s)$ can be formed by first writing

$$\frac{C(s)}{R(s)} = \frac{p(s)}{n(s)} \frac{G(s)}{1 + G(s)} \frac{p(s)}{p(s) + q(s)}$$

Note that the poles of the open-loop system are the zeros of $n(s) - p(s)$; therefore, if the real part of $n(s)$ is plotted on the same graph with the real part of $p(s)$, the roots of the open-loop system are the abscissas of the intersections of the two curves.

Note that for most systems the sum of the zeros of $q(s)$ is identical to the sum of the poles of $C(s)/R(s)$ (*i.e.*, the zeros of $n(s)$).

17.4.3.4. *Synthesis of Compensation Networks.* The desired system response is obtained by using as many poles and zeros of the controlled system as possible and canceling the others approximately with a compensation network.

17.4.4. Optimum Transient Response Behavior for Torque-Saturated Systems. Two of the torque-saturated curves pass through the origin (see Sec. 17.3.3.2), one for maximum positive torque, the other for maximum negative torque. These parabolas form the optimum switching curves (optimum implying the curve which takes the least time to reach the origin). Therefore, the fastest system will start heading toward the switching curve with maximum torque; when the switching curve is reached, the maximum opposite torque will be applied; this will take the system along the switching curve, to the origin. For small errors, or operation near the switching curve, a practical system based on this technique does not work too well. Therefore, a linear mode might be selected in these regions. The switching curve can be established using nonlinear elements.

17.4.5. Miscellaneous Comments on Design and Compensation

CRUDE ANALYSIS. Since in a closed-loop system

$$\frac{C(s)}{R(s)} = \frac{K_1 G_1(s)}{1 + K G_1(s) K_2 G_2(s)}$$

where $K_1 G_1(s)$ is the forward loop transfer function, and $K_2 G_2$ is the feedback loop transfer function,

$$\frac{C(s)}{R(s)} \approx \frac{1}{K_2 G_2(s)}$$

for frequencies where $K_1G_1(s)K_2G_2(s) >> 1$ and

$$\frac{C(s)}{R(s)} \approx K_1G_1(s)$$

for frequencies where $K_1G_1(s)K_2G_2(s) << 1$.

Inner Loop Gain Distribution. In two-loop systems, the inner loop gain should be distributed primarily in the forward loop (K_1), thus reducing the gain required in the outer loop.

Multiple-Loop Analysis and Synthesis: One-Third Rule. Generally, the response of some loops will be made faster than others, in order to maximize inner loop response, *e.g.*, a loop for stabilization in a track system. The faster loop will usually exhibit a peaking at its closed-loop frequency. Experience shows that the outer loop should have a closed-loop frequency no higher than one-third that of the inner stabilization loop.

Serial or Sequential Analysis. If the crossover frequencies of individual loops are far enough apart, the inner loops can be analyzed independently for stability and response, and the result used in the analysis of the outer loop.

Parallel Analysis. All loops having a common element are considered parallel. In Fig. 17-61 the common element is A, and thus the parallel loops are AB (loop 1); AD (loop 2); and ACE (loop 3). Then the output at any frequency is roughly that defined by the loop with the greatest gain at that frequency (or time). When considering the stability of the system, only the loop which predominates in that frequency domain must be considered.

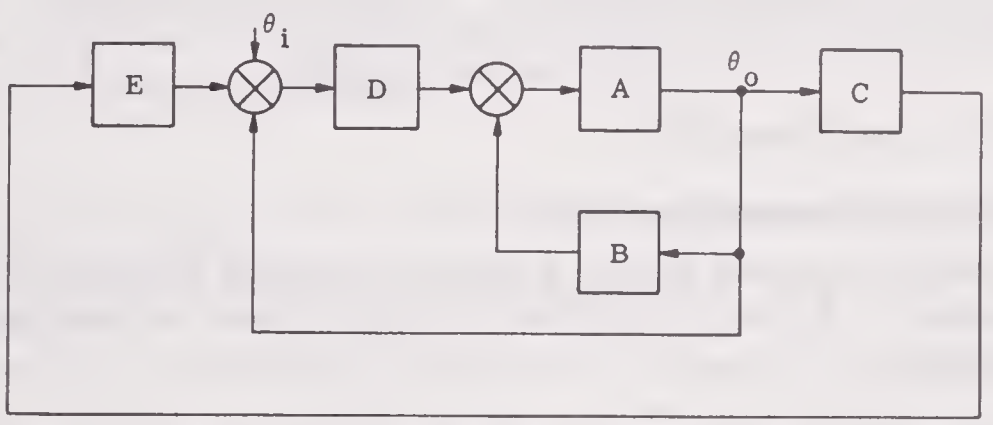


Fig. 17-61. Block diagram for parallel analysis.

17.4.6. Statistical Design

17.4.6.1. *Introduction.* Statistical design deals with the design of systems in which it is assumed that the input signals are from ergodic, stationary, random processes.

The error criterion most generally chosen is that the mean square error be reduced. This criterion places heavy emphasis on large errors and, as a consequence, often leads to a system design of low relative stability. It is, however, very convenient to handle mathematically and in many applications yields adequate results.

In designing a system statistically, it is often useful to know the probability distribution function of the random process. This can be found by determining what percentage of time a signal exceeds various amplitudes. Differentiating the result yields the probability distribution. This procedure can, of course, be carried out graphically.

If the mean-squared error criterion is chosen, signals are adequately described by correlation functions. These functions are defined by

$$\phi_{ij}(\tau) = \lim_{t \to \infty} \frac{1}{2t} \int_{-t}^{t} f_i(t) f_j(t + \tau)\, dt$$

If $i = j$, the result is the autocorrelation function; if $i \neq j$, the result is the cross-correlation function. The correlation function can be found graphically by sampling the function every α seconds and shifting the time axis to generate $f(t+\tau)$. Multiplying and averaging the results yields the desired function. The major characteristics of these functions are as follows:

(1) $\phi_{11}(\tau) = \phi_{11}(-\tau)$

(2) $\phi_{11}(0)$ is the average power

(3) $\phi_{11}(\infty) \leq |\phi_{11}(\tau)| \leq \phi_{11}(0)$

(4) If $f_1(t)$ contains periodic components, $\phi_{11}(\tau)$ contains components of the same period.

(5) If $f_1(t)$ contains a dc component, so does $\phi_{11}(\tau)$

(6) $\phi_{11}(\tau)$ contains no phase information

(7) If (4) and (5) do not hold, $\phi_{11}(\tau) \rightarrow 0$ as $\tau \rightarrow \infty$

(8) $\phi_{11}(\tau)$ is the sum of the autocorrelation function of the individual frequencies which make up $f_1(t)$

(9) $\phi_{11}(\tau)$ corresponds to an infinite number of time functions

(10) If $f_1(t) = f_a(t) + f_b(t)$, then

$$\phi_{11}(\tau) = \phi_{aa}(\tau) + \phi_{ab}(\tau) + \phi_{ba}(\tau) + \phi_{bb}(\tau)$$

(11) $\phi_{ab}(\tau) = \phi_{ba}(-\tau)$

(12) If $f_a(t)$ and $f_b(t)$ are uncorrelated (e.g., signal and noise), then

$$\phi_{11}(\tau) = \phi_{aa}(\tau) + \phi_{bb}(\tau)$$

(13) $\phi_{ab}(\tau) \leq [\phi_{aa}(0)\,\phi_{bb}(0)]^{1/2}$

where $\langle\ \rangle$ indicates the ensemble average.

The power-density spectrum $\Phi_{jj}(j\omega)$ is another important descriptor of signal and/or noise characteristics. $\Phi_{jj}(j\omega)$ can be found by using a spectrum analyzer, and has the following properties:

(1) $\Phi_{11}(j\omega) = F\{\phi_{11}(\tau)\} = \int_{\infty}^{\infty} \phi_{11}(\tau) e^{-j\omega\tau}\, d\tau$

(2) $\Phi_{11}(j\omega)$ has no phase information

(3) $\Phi_{11}(j\omega)$ does not correspond uniquely to a time function

(4) $\Phi_{11}(j\omega)$ is real

(5) $\Phi_{11}(j\omega) = \Phi_{11}(-j\omega)$

(6) $\Phi_{11}(j\omega) > 0$

(7) $\Phi_{11}(j\omega)$ contains impulses at frequencies contained in a periodic input.

Physical realizability of a system $G(s)$ can be determined as follows:

(1) $|G(j\omega)|$ is realizable if the Paley-Wiener Criterion is satisfied:

$$\int_0^{\infty} \frac{|\log|G_2(j\omega)||}{1+\omega^2}\, d\omega < \infty$$

(2) $g(t)$ is realizable if $g(t) = 0$ for $t < 0$ and $g(t) \rightarrow 0$ as $t \rightarrow \infty$

(3) If $G(s)$ is a ratio of polynomials, it is realizable if all the poles are in the left-half plane (but excluding ∞ and the $j\omega$ axis).

17.4.6.2. *Design Principles.* If physical realizability of the system is not considered, the optimum transfer function $G_{opt}(j\omega)$ for minimizing the mean-square error can be found from

$$G_{opt}(j\omega) = \frac{\Phi_{ss}(j\omega)}{\Phi_{ss}(j\omega) + \Phi_{nn}(j\omega)} G_d(j\omega)$$

where $\Phi_{ss}(j\omega)$ = power density spectrum of signal

$\Phi_{nn}(j\omega)$ = power density spectrum of noise

$G_d(j\omega)$ = desired system transfer function

If physical realizability is a consideration, the optimum transfer function $G_{opt}(s)$ is given by

$$G_{opt}(s) = \left[\frac{\Phi_{ss}(s)}{\Phi_{ii}^{-}(s)}\right]^{+} \frac{1}{\Phi_{ii}^{+}(s)} G_d(s)$$

where

$$\Phi_{ii} = \Phi_{ss} + \Phi_{nn} = \Phi_{ii} + (s)\Phi_{ii}^{-}(s)$$

and where the superscript + means all critical frequencies in the left-half plane, and the superscript − means all critical frequencies in the right-half plane.

It should be noted that this design is based on the assumptions that the mean-square error criterion is adequate, that signal and noise are stationary time series, and that a linear, series compensation is desired. The problems associated with this method are that $\Phi_{ss}(s)$ is difficult to obtain, the derived system often has very light damping and is sensitive to parameter changes.

References

1. H. James, N. Nichols, and R. Phillips: "Theory of Servomechanisms," McGraw-Hill Book Company, Inc., New York, 1947.
2. J. Truxal: "Automatic Feedback Control System Synthesis," McGraw-Hill Book Company, Inc., New York, 1955.
3. J. Tou: "Digital and Sampled-Data Control Systems," McGraw-Hill Book Company, Inc., New York, 1959.
4. M. Gardner and J. Barnes: "Transients in Linear Systems," vol. 1, John Wiley and Sons, Inc., New York, 1954.
5. J. Truxal: "Control Engineers' Handbook," McGraw-Hill Book Company, Inc., New York, 1958.
6. G. Thaler and R. Brown: "Servomechanism Analysis," McGraw-Hill Book Company, Inc., New York, 1953.
7. Z. Rekasius and J. Gibson: *IRE, Trans. Auto. Control* **7,** 3 (1962).

Bibliography

Ahrendt, W., and J. Taplin: "Automatic Feedback Control," McGraw-Hill Book Company, Inc., New York, 1951.

Andronow, A. A., and C. E. Chaikin: "Theory of Oscillations," Princeton University Press, Princeton, New Jersey, 1949.

Bellman, R.: "Stability Theory of Differential Equations," McGraw-Hill Book Company, Inc., New York, 1953.

Brown, G., and D. Campbell: "Principles of Servomechanisms," John Wiley and Sons, Inc., New York, 1948.

Chestnut, H., and R. Mayer: "Servomechanisms and Regulating System Design," vol. II, John Wiley and Sons, Inc., New York, 1955.

Evans, W. R.: "Control-System Dynamics," McGraw-Hill Book Co., Inc., New York, 1954.

Guillemin, E. A.: "Communication Networks," vol. II, John Wiley and Sons, Inc., New York, 1935.

Guillemin, E. A.: "The Mathematics of Circuit Analysis," John Wiley and Sons, Inc., New York, 1949.

Guillemin, E. A.: *Proc. IRE Natl. Electron. Conf.* **9,** 1953 (1954).

Grief, H. D.: *Trans. AIEE* **72,** 243 (1953).

Kazda, L. F.. "Proceedings of Workshop Session in Lyapunov's Second Method" (The University of Michigan, Ann Arbor, Michigan, 1962).
Liénard, A.: Revue Générale de l'électricité **23,** 901 (1928).
Mulligan, J. H.: *Proc. IRE* **37,** 516 (1949).
Ragazzini, J. R., and L. A. Zadek: *Trans. AIEE* **71,** 225 (1952).
"Reference Data for Radio Engineers," 4th Edition, International Telephone and Telegraph Corp., New York, July 1957.
Seamans, R. L., Jr., B. P. Plasingame, and R. C. Clentson: *J. Aeron. Sci.* **17,** 22 (1950).
Valley, G. E., and H. Wallman: "Vacuum Tube Amplifiers," McGraw-Hill Book Company, Inc., New York, 1948.
Zadeh, L. A.: *Proc. IRE* **38,** 291 (1950).

Chapter 18

SYSTEM DESIGN

Sol Shapiro

General Precision, Inc.

William L. Wolfe

The University of Michigan

CONTENTS

18.1. Design Procedures; Generalities 730
18.1.1. Example of Procedure for Gross Analysis 730
18.1.2. Sensitivity Analysis 731
18.2. Equations for Sensitivity Calculations. 731
18.3. Optical System 732
18.4. Scanning Dynamics 734
18.5. Scanning Techniques. 735
18.5.1. Rotating Wedges 735
18.5.2. Other Scanning Methods 736
18.6. Example of Design for a Search System 737
18.6.1. Requirements Analysis. 737
18.6.2. Design Analysis 737
18.6.3. Definition of Primary System Parameters 740
18.7. Tracking System Design 743
18.7.1. Conception 744
18.7.2. Sample System 746
18.8. Mapping Systems 750
18.8.1. Requirements 750
18.8.2. Basic Sensitivity Equation 751
18.8.3. Background-Limited D^*-Q Trade-Off 753
18.8.4. Sample System 754

18. System Design

18.1. Design Procedures; Generalities

The design of infrared systems requires the employment of many engineering approximations, compromises, and judgments about when to approximate and how to compromise. This chapter provides some hints and some equations for guidance. It should be understood that not all of them necessarily apply in all cases.

The formation of the problem usually involves specifications or at least judgments about the following; this constitutes requirements analysis:

1. Function
2. Volume of coverage
3. Volume coverage rate
4. Desired resolution
5. Range
6. Target type
7. Background
8. Operating altitude
9. Desired output form
10. Size, weight, reliability, etc.

The design analysis steps can often be thought of as:

1. Gross implementation
2. Sensitivity analysis
3. Scanning or tracking analysis
4. Output analysis

Sensitivity analysis and scanning analysis are intimately related.

18.1.1. Example of Procedure for Gross Analysis.

1. Make an estimate of entrance-aperture flux density from the target radiation based on blackbody radiation and some chosen emissivity – or a better quick estimate if available.
2. Estimate background fluctuations in about the same way, choosing "good" values for emissivity and temperature changes and an instantaneous field of about 1 mil, or one matching the target.
3. Calculate dwell times necessary from instantaneous field, total field, and required frame rate.
4. See if these are compatible with existing optics, detectors, rotation rates, reticle spacings, etc.
5. Either refine the design of item 4, or "bend" some of the constraints, or change some assumed system parameters drastically.

18.1.2. Sensitivity Analysis. Stated in its simplest terms the problem is to calculate the ratio of the target signal to the total noise; the total noise can usually be calculated as the sum of the background noise and the system noise. Thus:

1. Calculate the spectral, temporal, spatial characteristics of the target radiance.
2. Calculate the attenuation by the atmosphere of this signal.
3. Repeat items 1 and 2 for the background.
4. Assume an optical system (usually $f/1$, 50% efficiency) and choose an appropriate detector. Calculate system noise and aperture size necessary for detection.
5. If step 4 is successful, stop; if it is not, choose other combinations that are successful – possibly changing the number of detectors to decrease the noise bandwidth, etc.

After this, specific parameters and special problems must be considered.

18.2. Equations for Sensitivity Calculation

Many workers prefer to calculate the noise equivalent input on a step-by-step basis. For those who do not, the equations for a field-mapping application (target smaller than instantaneous field of view) and for tracking (target filling the field and target less than full field) are given below.

Mapping sensitivity

$$\Delta H \geq \frac{4k\mathcal{F}^2}{\tau_0 \sqrt{AA_d}} \int \mathrm{NEP}'\, df$$

Target-detection sensitivity

Target > field

$$H(T, \epsilon, r) \geq (K+1)H(T_1, \epsilon_1, r) - KH(T_0, \epsilon_o, r)$$

$$H(T_1, \epsilon_1, r) - H(T_0, \epsilon_o, r) \geq k\mathrm{NEI}$$

Target < field

$$\Delta \mathrm{H}(T, \epsilon, r) \geq \sigma^2 r^2 (k+1) H(T_1, \epsilon_1, r) - [\sigma^2 r^2 (k+1) - A]\, H(T_0, \epsilon_o, r)$$

NEP (noise equivalent power) for scanning systems of this type using a single detector with detectivity D^* is

$$\mathrm{NEP} = \frac{FD\ \Omega^{1/2}}{D^* T^{1/2}}$$

NEI (noise equivalent input) is

$$\mathrm{NEI} = \frac{\mathrm{NEP}}{(\pi D^2/4)\ \epsilon_o \gamma} = \frac{4F\ \Omega^{1/2}}{\pi \epsilon_o \gamma D D^* T^{1/2}}$$

For a system with N_0 equal detectors, the solid angle scanned by each detector can be reduced by the number of detectors; thus

$$\mathrm{NEI} = \frac{4F}{\pi \tau_0 \gamma D D^*} \frac{\Omega^{1/2}}{\sqrt{N_0 T}}$$

In the equations above,

A = target area
B = Planck function
D = diameter of entrance pupil
D^* = detectivity
F = focal length
$\mathfrak{F}$ = focal ratio
H = irradiance
k = confidence constant
K = ratio of target signal to background signal
r = range
T = temperature
γ = system response to target radiation
Γ = detector relative response
ϵ = emissivity
σ = Stefan-Boltzmann constant
τ_a = atmospheric transmission
τ_o = optics transmission
Ω = solid angle
Subscripts 1 and 0 = specific values

18.3. Optical System

The following functions are usually most important in optical system design generalities:

1. Sensitivity: use large relative aperture (f/no), large aperture, high efficiency.
2. Background discrimination: make instantaneous field equal to target subtense.
3. Coverage: large enough to cover required area, but small to reduce scan rates.
4. Simplicity: keep simple, light, inexpensive, etc.

To achieve sensitivity, a field lens and reimaging technique can sometimes be used. A lens is placed about where the primary image is formed. It is made large enough to gather all the rays, and usually images the entrance pupil onto the detectors. The optics no longer image the target, but they uniformly irradiate the detector (see Sec. 10.2.4). The effective f/no of the system is usually the angle of convergence on the detector but sometimes includes obscuration effects. Primary optics are chosen from among the Schmidt, Cassegrain, Bouwers, and Newtonian telescopes. Others are sometimes used, and perhaps someday lenses may be used. Depending upon the application, the entrance pupil or the image plane is imaged onto the detector.

The equations for imaging the image are:

$$L = 2F \tan \Theta/2$$

$$d \approx \frac{Do}{F} + L$$

$$\theta = \tan^{-1} L/2o$$

$$i = \frac{of}{o - f}$$

where Θ = field angle

θ = secondary field angle

F = primary focal length

f = secondary focal length

D = primary diameter

d = secondary diameter

L = maximum linear dimension of image plane

o = object distance for secondary lens

i = image distance for secondary lens

A typical problem can be stated as follows: for a primary objective of diameter D, image distance F (focal length for infinitely distant object points), and total field angle Θ, design a field lens which collects all the radiation and which refocuses the image at some given distance s behind the primary focus. The designer must then decide how big to make the lens, and what focal length and what field of view to use.

The equations can be normalized with respect to L. The geometry is given in Fig. 18-1. The equations are

$$\frac{d}{L} \approx \frac{D}{F}\frac{o}{L} + 1 = \frac{1}{\mathcal{F}}\frac{o}{L} + 1$$

$$\theta = \text{arc cot } 2o/L$$

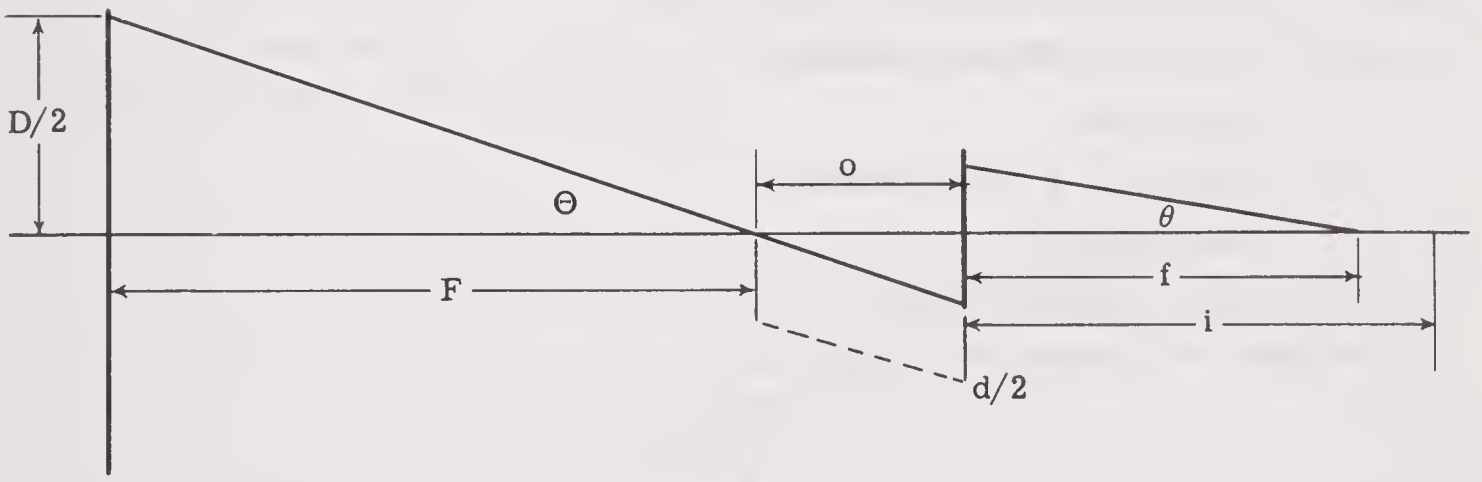


FIG. 18-1. Relay optical system.

A second type of problem calls for imaging the entrance pupil (usually the same as the aperture stop and the primary) onto the image plane by placing a lens at the primary image. Then,

$$o = F + s$$

$$d = \frac{D}{F}s + L$$

where s = distance of lens behind primary image.

The angular field coverage is not important; it is only necessary to collect all the energy and "smear" it over the surface of the detector.

18.4. Scanning Dynamics

The problem is to choose an optimum kind of scan pattern and to determine the realism of a chosen instantaneous field, total field, frame rate, detector number, etc.

The required dwell time is

$$t = \omega T/\Omega$$

where ω = solid angle of instantaneous field

Ω = total solid angle

t = dwell time per detector element

T = allowable field time

If a single detector is used, it must have a time constant τ given by

$$\tau = t/g = \omega T/g\Omega$$

where g = number of time constants per dwell time (usually 1 to 3). The following equations give prism rotation rates and instantaneous fields of view for a mapping system. (The meanings of r, θ, a, and b are made clear in Fig. 18-2.)

$$s = \sqrt{(2\pi gnp)^{-1}(v/h)\ (1/\tau)}$$

$$\sigma = \sqrt{(2\pi n/gp)(v/h)\tau}$$

$$b \approx h\sigma \sec^2 \theta$$

$$a = \sigma h \sec \theta$$

where σ = instantaneous angular field of view

p = number of detector elements

s = prism rate

$1/n$ = fraction of a circle scanned

τ = time constant

v/h = velocity-to-height ratio

If an array of N detectors is used, then

$$\tau = \frac{N}{g}\frac{\omega}{\Omega}T$$

Usually a scanning system has less than 100% duty cycle. Several consequences are possible singly or in combination.

1. Less field is covered.
2. The field is covered less frequently.
3. The scan is speeded.
4. More detectors are added or other design changes are made.

If the duty cycle is increased, the dwell time can be increased:

$$\tau = D_c\frac{N}{g}\frac{\omega}{\Omega}T$$

where D_c = fractional duty cycle.

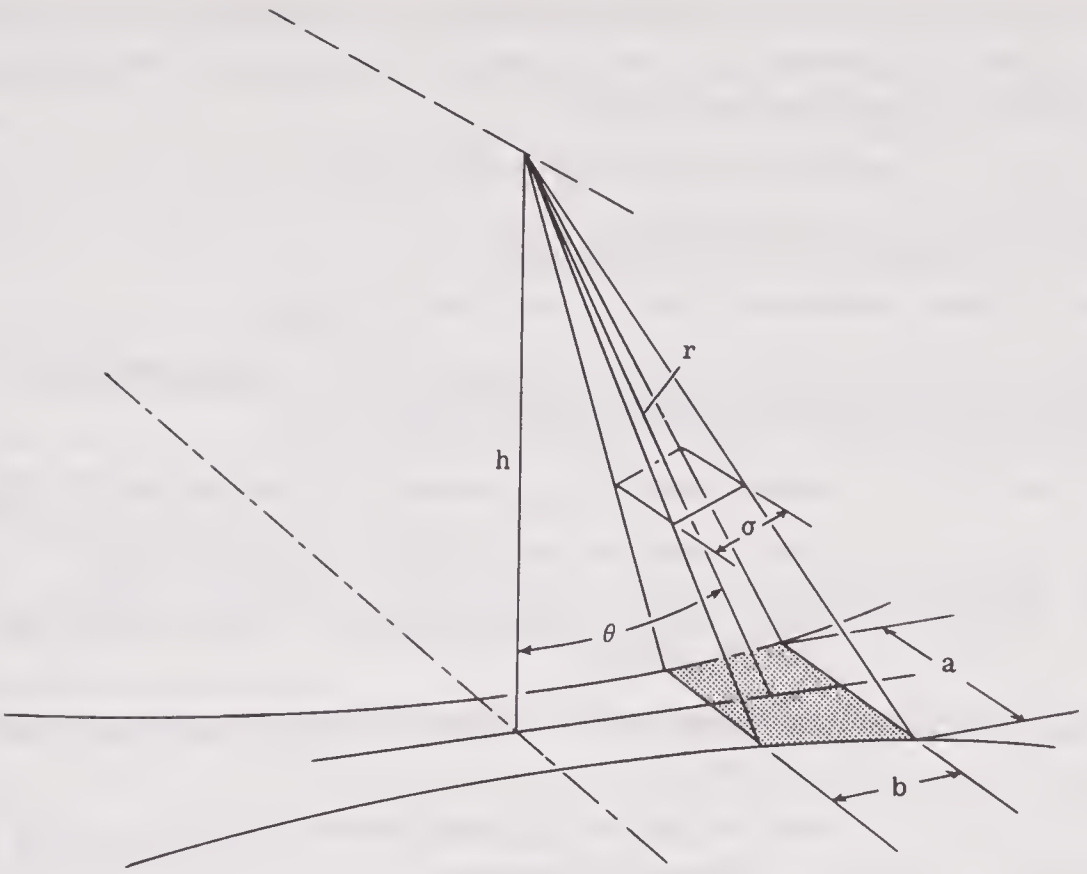


FIG. 18-2. Scanning geometry.

The required bandwidth is related to the time of a single look at the instantaneous field of view. To a good approximation,

$$\Delta f = 1/2\tau$$

18.5. Scanning Techniques

18.5.1. Rotating Wedges. A single wedge deviates the ray by the usual deviation equation for a prism. The equations are

$$\delta = \theta_1 + \theta_2 + \alpha$$

$$\sin \theta_2 = \frac{n'}{n} \sin \left\{ \pi - \left[\alpha + \arcsin \left(\frac{n}{n'} \sin \theta_1 \right) \right] \right\}$$

$$= \frac{n'}{n} \sin \left[\alpha + \arcsin \left(\frac{n}{n'} \sin \theta_1 \right) \right]$$

In the special case of normal incidence,

$$\sin \theta_2 = \frac{n'}{n} \sin \alpha$$

where n' = prism material index

n = environment index

θ_i = incident or exit angle

α = prism angle

δ = deviation angle

The refraction takes place in a single plane so that a ray is bent only in one direction. Assume that the direction is the x axis; then θ_2 is θ_x. As the prism is rotated about the z axis, the angular direction of the ray changes. The resultant angle can be thought

of as the combination of two vectors determined by the position of the prism and its refraction. When two prisms are used in series the deviations are combined. Two angular rotations are summed vectorially:

$$\theta = \theta_1 + \theta_2 = |\theta_1|\ e^{j(\omega_1 t + \delta_1)} + |\theta_2|\ e^{j(\omega_2 t + \delta_2)}$$

where ω_i is the rotation rate of the ith wedge

δ_i is the phase reference to the x-axis of the ith wedge

By consideration of the relative rates of rotation and the phase angles of two equal-angle prisms several useful scan patterns can be generated:

$\omega_1 = \omega_2, \quad \delta_1 = \delta_2 \quad \Rightarrow$ a circle is generated as if there were a bigger prism

$\omega_1 = -\omega_2, \quad \delta_1 - \delta_2 = \pi, \Rightarrow$ a point results as if there were no prism

$\omega_1 = \omega_2, \quad \delta_1 \pm \delta_2 = k \quad \Rightarrow$ a circle is generated with phase delay and reduced amplitude

$\omega_1 = \pm k\omega_2, k =$ constant $\Rightarrow$ loops within a circle result; the greater the k the more loops

$\omega_1 \neq \omega_2, \quad \Rightarrow$ a spiral is generated with loops decreasing as the inequality increases

18.5.2. Other Scanning Methods.

Of the many types of scan pattern, the most useful ones are:

1. Circular
2. Palmer
3. Spiral
4. Rosette
5. Hypocycloid
6. Raster

The circular scan can be generated by a rotating off-axis mirror, lens, prism, etc. Either the primary, secondary, or a folding flat mirror or a wedge can be used. The radius of the circle depends upon the angle generated and the distance from the generator to the image plane. The angle of ray inclination is about equal to the angle of inclination of a lens, usually twice the angle of inclination for a mirror and equal to the deviation angle for a prism or plate. (These deviation considerations apply to the other scans as well.)

The Palmer scan can be generated by translation of the circular scan; geometrical considerations provide data on overlap.

Spiral scans can be generated by a second motion superimposed on the rotating elements providing circular scan. The motion is a variable tilt angle.

Rosettes and hypocycloids are generated by superimposing two circular scanning motions. Hypocycloids are generated by the superposition of two circular motions. If the slowly moving element has a much larger deviation, the pattern is almost a small sinusoidal excursion on a circle. The number of loops in the pattern depends upon the relative speeds of the two motions, and the depth of the "petals" depends on the relative angular deviations.

Rasters are usually combinations of two orthogonal linear motions.

A raster scan is usually to be chosen for search and imaging work (particularly when no prior position information is available); rosettes and hypocycloids are useful for tracking.

18.6. Example of Design for a Search System

This discussion of the design of a representative search system is divided into (*a*) requirements analysis; (*b*) design analysis; (*c*) definition of primary system parameters; and (*d*) specific problems. It is an outgrowth of one author's specific experience and as such represents a combination of the several system designs in which he was involved. The problem is to design a ballistic-missile detection system operating from an aircraft.

18.6.1. Requirements Analysis.

18.6.1.1. *Target Radiation.* The assumed target for this discussion is a solid-fueled missile; this can be crudely approximated by a 1000°C blackbody that provides a total radiant intensity of 200,000 w sr^{-1}. The altitude-time profile of the target is parabolic, reaching 300,000 ft in 100 sec.

18.6.1.2. *Problem Geometry.* The range at which ballistic missile detection is generally desired requires consideration of earth curvature to define the limiting conditions of detection range.

To find the maximum sighting range to the target at any given altitude, one must calculate the sum of the target distance and sensor distance to a line tangent to the effective horizon. The altitude of the effective horizon above the surface of the earth is a function of cloud cover and transmission. One must find a proper minimum horizon altitude h_0 for his spectral region.

In this sample problem, detection of the missile from an aircraft at 40,000 ft with the horizon at 30,000 ft will be assumed. The result of the calculations described above is to show the length of time the target is in view before burnout as a function of target-sensor range. Such a curve would show the time available for detection as a function of problem geometry. This geometry-directed range is approximately 480 n mi.

18.6.1.3. *Atmospheric Transmission, Radiation, and Spatial Filtering.* The transmission of the atmosphere is required to determine the irradiance from the target in the spectral region of detector response. A relatively crude calculation of transmission in the detector bandpass is normally adequate in defining initial system design parameters. For purposes of this example, two spectral bands will be considered: 1.8 to 2.7 μ (lead sulfide) and 3 to 5 μ (indium antimonide). Net effective integrated transmission of the bands will be taken as 30% over each band. In a real problem, if selection of a wavelength band of operation is important, detailed analysis of the transmission supported by repeatable missile radiation analysis is desirable.

The radiation associated with the atmosphere and discrimination against it in favor of a real target is often the most significant atmospheric effect in infrared systems. In the current analysis, the problem would be that of thunderheads and cloud tops presenting possible false targets. The most direct approach in achieving this discrimination is to recognize that the real target is far more intense per solid angle of its subtense than the background. Thus the search system's instantaneous field of view is made as small as feasible. The background radiation in the field and gradient effects (cloud edges) are then minimized. The worst background level from sun reflection may be approximated by 10^{-2} w cm^{-2} sr^{-1} with a 6000°K spectrum which when viewed through the approximate 30% atmospheric transmission yields 3×10^{-3} w cm^{-2} (of aperture) sr^{-1}. From self-emission through the 30% atmosphere, the irradiance at the entrance pupil is also approximately 3×10^{-3} w cm^{-2} sr^{-1} with a 0°C spectrum.

18.6.2. Design Analysis.

18.6.2.1. *Mode of Operation.* The purpose of the sample search set is detection, not track-while-scan or surveillance. When it has provided positive identification

of a target and sufficiently accurate angular information for other equipment to operate on the target, its function is complete.

A detection is defined if a target-like signal is seen on two successive scans. The two detections are necessary because of the existence of nongaussian "noise" such as power-supply transients in a system. Such a two-detection system is better than a two-out-of-three system because (neglecting the effects of nongaussian noise) it can be shown that designs for three detections have reduced the signal-to-noise ratio by a factor of $\sqrt{2/3}$ (82%). The higher signal-to-noise ratio achieved in the two-detection system more than offsets its lower probability of detection.

The azimuth coverage will be 360°. The range will be 90% of the geometry-limited range, or 433 n mi. The frame time is defined by the requirement for two full scans in the time the target moves from the altitude corresponding to 90% of range to burnout. This gives a viewing time of 6 sec and a resulting frame time of 3 sec. Selection of an elevation of field of view in this kind of detection system involves the problem of minimum range of operation of the "defense screen."

Now one plots the altitude of the horizon line against range from the defense aircraft and the missile altitude two frame times after crossing the horizon altitude line. From this one computes or plots the elevation angle coverage necessary to obtain detection as a function of minimum range operation. For purposes of this problem, a minimum range of 40 n mi will be assumed, and for an assumed missile velocity profile at 40 n mi, the elevation coverage required is 3°.

18.6.2.2. *Wavelength of Operation.* Two wavelength regions 1.8 to 2.7 μ and 3 to 5 μ, will at least initially be carried as parallel analyses. To a first approximation sensitivity of the two systems to the 1000°C blackbody target will be sufficiently close together so that the choice of wavelength must be based on other factors. These factors include:

Background discrimination: Discrimination at 3-5 μ with similar techniques is better than that achievable at 1.8 to 2.7 μ.

Complexity and cost: Operation at 3 to 5 μ requires cooled detectors, with attendant problems of complexity and cost.

Detector time constant: In order to achieve background discrimination, small instantaneous field of view and the resultant high-speed scan is desirable. This requires the use of a short-time-constant detector which would be more readily available at 3 to 5 μ.

The trade-off is then one of cost and complexity with reduced background-discrimination performance.

18.6.2.3. *Sensitivity.* Effective target radiation from the 1000°C source radiating 200,000 w sr^{-1} through a 30% transmission atmosphere between 1.8 and 2.7 μ is:

$$S_{\Delta\lambda} = (S_{total})(\%\text{ in band})(\text{atm transmission})$$

$$S_{1.8\text{-}2.7} = (200{,}000)(0.25)(0.3)$$

$$= 15{,}000 \text{ w sr}^{-1}$$

In the 3-5-μ band,

$$S_{3\text{-}5} = (200{,}000)(0.32)(0.30)$$

$$= 19{,}200 \text{ w sr}^{-1}$$

At the maximum design detection range of 433 n mi,

$$H_{1.8\text{-}2.7} = \frac{S_{1.8\text{-}2.7}}{R_{max}{}^2}$$

$$= \frac{15{,}000}{(433)^2(6080)^2(144)(6.45)}$$

$$= 2.3 \times 10^{-12} \text{ w cm}^{-2}$$

$$H_{3\text{-}5} = \frac{19{,}200}{6.46 \times 10^{15}}$$

$$= 3.0 \times 10^{-12} \text{ w cm}^{-2}$$

If background radiation in the instantaneous field of view is to be about that of the target, the required field of view to discriminate against sun scattering at 1.8 to 2.7 μ is

$$\omega = \frac{H_{target}}{N_{sun\ reflections}} = \frac{H_{target}}{(\text{sun radiation})\,(\text{reflection})\,(\%\text{ in band})}$$

$$= \frac{2.3 \times 10^{-12}}{(3 \times 10^{-3})(0.05)}$$

$$= 1.5 \times 10^{-8} \text{ sr}$$

For the 3-5-μ band, the angular field of view needed for discrimination against sun reflections is

$$\omega = \frac{3 \times 10^{-12}}{(3 \times 10^{-3})(0.015)}$$

$$= 6.7 \times 10^{-8} \text{ sr}$$

The self-emission (0°C source, 3×10^{-3} w cm^{-2} sr^{-1}) is less than 1% of total in the 3-5-μ band and so is less of a problem than sun reflection. No consideration has been given to the practical problem of small-detector fabrication.

It is now possible to define the dwell time of a target in the instantaneous field of view consistent with optimized background discrimination as a function of the number of detectors in the search system, for the moment again without consideration of the practical problem of fabrication of miniscule detectors.

$$t = \frac{N\omega}{\Omega} T$$

$$t_{(1.8\text{-}2.7)} = \frac{3(1.5 \times 10^{-8})N}{(2\pi)(3/57.3)}$$

$$= (1.4 \times 10^{-7})N \text{ sec}$$

Similarly,

$$t_{(3\text{-}5)} = \frac{3(6.7 \times 10^{-8})N}{(2\pi)(3/57.3)}$$

$$= (6.0 \times 10^{-7})N \text{ sec}$$

At this point it appears that, even for $N = 100$, if the practical problems of small-detector fabrication are solved so that spatial filtering to very small angles is achieved, the use of lead sulfide in the 1.8-2.7-μ band must be eliminated on the basis of the time constant. On the other hand lead sulfide can be considered if, instead of an individual detector field of view of 0.015 to 0.067 μ sr^{-1} (linear square of 0.1 to 0.3 mrad), the individual field is 1 μ steradian or more with background rejection by other techniques.

To define system sensitivity requirement, the signal-to-noise ratio required to provide a specified probability of detection consistent with a specified false alarm rate must be defined.

It will be assumed that one false alarm per 10^7 sec is tolerable. A false alarm is defined as repetition of a false signal on two successive scans within an area of 100 (10×10) resolution elements. It is assumed that processing in a computer will provide this logic. With one scan every 3 sec, in 10^7 sec, there are 3×10^6 scans. There are also essentially 3×10^6 successive pairs of frames (1 and 2, 2 and 3, etc.). The probability of detection per successive scan pair required is then about $(3 \times 10^6)^{-1}$, or about 3×10^{-7}. Any single scan may be paired with the preceding or the following scan, so that single scan detection probability must be

$$2P_{1\ scan}{}^2 = 3 \times 10^{-7}$$

$$P_{1\ scan} = 4 \times 10^{-4}$$

To define the required operating threshold relative to rms noise: the number of noise peaks per second is:

$$\dot{n}_p = 0.775\ \Delta f$$

also,

$$f_{\text{upper}} = 1/2t$$

thus

$$\dot{n}_p = 0.39/t$$

The number of noise peaks in 100 dwell times is then 39. The probability of any single peak being above threshold must then be:

$$4 \times 10^{-4}/39 = 10 \times 10^{-6}$$

From gaussian probability,

$$P(I) < y\ \sqrt{\psi_0} = 1.5\ e^{-y^2/2}$$

the threshold y is computed as 4.9 times the rms noise.

18.6.2.4. *Probability of Detection and Required Signal-to-Noise Ratio.* For purposes of this problem, the probability of detection will be taken as 99.8% for the two scan situation. The single-scan probability of detection must then be $\sqrt{99.8\%}$, or 99.9%. With the threshold (ν) set at 4.9 times rms noise, the required signal-to-noise ratio for 99.9% single-scan probability of detection is 8.2. To achieve this signal-to-noise ratio on the target signal, the noise equivalent input of the system for operation in the 1.8-2.7-μ band is:

$$\text{NEI}_{1.8\text{-}2.7} = 2.3 \times 10^{-12}/8.2$$

$$= 2.7 \times 10^{-13}\ \text{w cm}^{-2}$$

Similarly,

$$\text{NEI}_{3\text{-}5} = 3 \times 10^{-12}/8.2$$

$$= 3.6 \times 10^{-13}\ \text{w cm}^{-2}$$

18.6.3. Definition of Primary System Parameters.

18.6.3.1. *Sensitivity Parameters.* It will be shown that, to a first order, the system may be defined by the expression

$$\text{NEI} = \frac{4\Omega^{1/2}}{\pi DD^*\ \epsilon_0\ \sqrt{NT}}$$

where NEI is noise equivalent input

Ω is solid field angle scanned per frame

D is optic diameter

D^* is detector detectivity

N is number of detectors

T is frame time

ϵ_0 is optical transmission

$$\text{NEI}_{1.8\text{-}2.7} = 2.7 \times 10^{-13} \text{ w cm}^{-2}$$

The values for inclusion in the equations are

$$\Omega = 2\pi(3/57.3) = 0.33 \text{ sr}$$

$$T = 3 \text{ sec}$$

$$D^* = 5 \times 10^{10} \text{ cm (w sec)}^{-1/2}$$

Then:

$$D\sqrt{N} = \frac{5.2\sqrt{0.33}}{5 \times 10^{10}\sqrt{3}(2.7 \times 10^{-13})}$$

$$= 130$$

For 3-5 μ and with $D^* = 5 \times 10^{10}$

$$D\sqrt{N} = 100$$

18.6.3.2. *Selection of Specific Values for D and N.* Selection of specific D and N to achieve the required sensitivity involves trading off size, weight, complexity, and reliability, and detector array availability as well as scan-pattern generation if the full 3° elevation field is not covered by the array (say to meet background rejection by filtering).

Neglecting spatial filtering for the moment and considering current detector array availability, a 10-in. aperture with 25 lead sulfide or 15 liquid-nitrogen-cooled indium antimonide detectors would seem to be a reasonable compromise. The fields of view and dwell time of square detectors are:

1.8-2.7 μ

$$\omega = \left(\frac{3}{25}\right)^2 = 0.014^{\circ 2} = 4.3 \ \mu \text{ sr}$$

$$T_0 = \frac{4.3}{0.015}(1.4 \times 10^{-7})(25) = 1.0 \text{ msec}$$

3-5 μ

$$\omega = \left(\frac{3}{15}\right)^2 = 0.04^{\circ 2} = 12 \ \mu \text{ sr}$$

$$T_0 = \frac{0.12}{0.067}(6.0 \times 10^{-7})(15) = 1.6 \text{ msec}$$

With the lead sulfide, the dwell time is essentially at the reasonable time-constant limit. The ratio of maximum or worst background power to target power in the field is 4.3/0.015 ~ 300. If pulse length or cloud edge slope discrimination can be used to decrease this ratio, the lead sulfide system may be operable.

The most promising system from a performance standpoint is indium antimonide where the ratio of worst background to signal detection is 12/0.067 ~ 180. However, since the dwell time of 1.6 msec is far from detector time constant limited, it may be possible to reduce the detector width to 0.3 mrad (0.003 in. in the 10-in. $f/1$ system) thereby giving a worst background signal only 15 times target radiation and offering better hope of pulse-length discrimination. This indium antimonide system with pulse-length discrimination would be the primary system.

18.6.3.3. *Other System Parameters.*

Optics. The 3° field-of-view, $f/1$ system requiring resolution of 0.2 mrad at 1.5° off axis is within current design capacity.

System Block Diagram. For the signal-processing gimbal drive and stabilization, see Fig. 18-3.

Cooling. A closed-loop liquid-nitrogen-transfer system can be developed for this application.

Signal Processing. The primary signal processing system uses separate solid-state preamplifiers for each detector and is within the state of the art for lead sulfide or indium antimonide. A hard agc loop referenced to a common buss is used to insure that the average noise level out of all detectors is the same. Saturation of the agc signal only slightly above noise peaks (say 10 to 11 times rms noise) insures that any strong signals will not seriously degrade channel sensitivity. The outputs of the

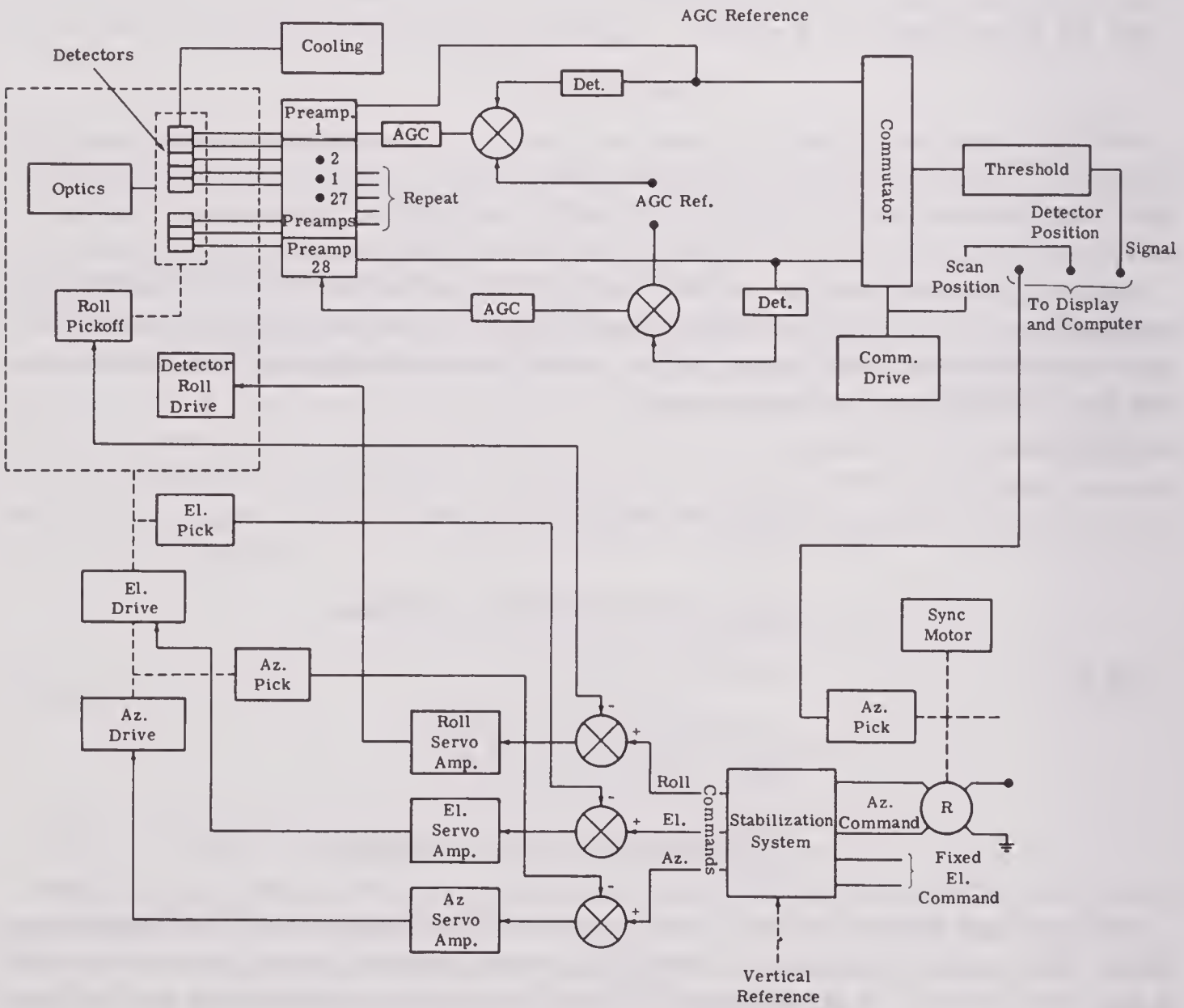


FIG. 18-3. Possible signal processing.

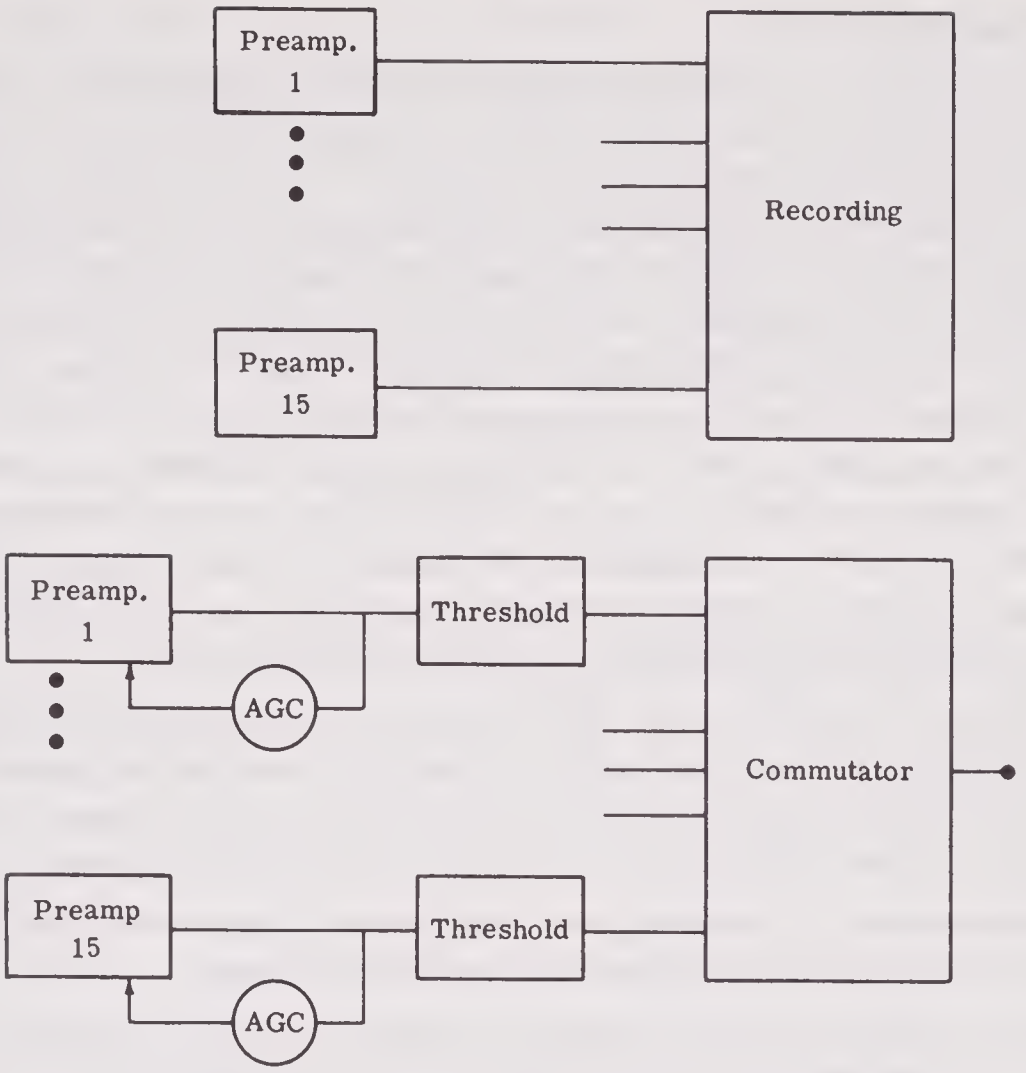


FIG. 18-4. Alternate signal processing.

detectors are commutated using diode logic matrices. A threshold set at 5.3 times rms noise is used at the output of the commutator. Signals above threshold are available for display and computer logic use.

Alternate signal processing is shown in Fig. 18-4, for recording only and for separate agc and threshold before commutation. Choice of threshold before or after commutation is a detail design decision.

Because of the continuous rotation of the scanner and the resultant need for slip rings, location of the commutator on the gimbals would be very desirable. The preamplifiers should be placed close to the detectors.

A continuous rotation-stabilized scan is to be used. This command is generated by a driven resolver. It is stabilized from a vertical reference system to provide azimuth and elevation gimbal position commands and detector roll stabilization. The scan will be parallel to the horizon to avoid crossing the high gradient of the horizon. Azimuth and elevation stabilization are conventional in search systems. The detector roll stabilization serves in large vehicle roll angles to maintain full field coverage. Because of the symmetry of the optics, they need not be roll stabilized.

Scan position pickoff will be generated from the azimuth program command. Servo slaving errors will be made small enough to make this an accurate indication of position. Elevation position from commutator drive will also be provided to computer and display.

18.7. Tracking System Design

In order to develop and discuss problems which are related to track system design, a specific problem, that of tracking aircraft from an aircraft-borne tracker will be used. Analogies to other combinations can be drawn by the system designer.

18.7.1. Conception.

18.7.1.1. *Acquisition.* In the conventional system, a search set (which may be integral with the tracker) will have defined the target position with some accuracy. This acquisition field of existing systems is about 2° to 4°. In general, also, the target will be at a relatively long range so that problem geometry cannot include very large angular rates of the line of sight. The problem of acquisition, then, is generally to bring a quasistationary target to the boresight of the tracker in a time which is short compared with the total problem time. Problem time in air-to-air intercept will actually be measured in tens of seconds so that acquisition time is 0.33 to 0.5 sec from initiation of the acquisition cycle. In a linear system, this defines a servo loop bandwidth whose response is about 0.1 sec or a closed-loop frequency of 10 rad/sec. Whether a linear system is used in acquisition is a function of overall system considerations including such factors as whether acquisition and track use the same error-sensing and servo-loop logic and the specifics of error-sensing techniques.

18.7.1.2. *Track.* It can be shown that in air-to-air intercept, once a target has been acquired, a second-order track-loop servo and stabilization against aircraft motion can keep track errors to less than 1/4°. Thus, the track field of view diameter may be made as small as 1/2°. Some advantages of using the small track field are:

1. Increased sensitivity, allowing continued track if target radiation is reduced by aspect or target use of radiation suppression techniques.
2. Background discrimination, including cases in which the target may move across severe backgrounds such as the horizon or bright cloud edges. The small field reduces the probability of severe background in the field and the intensity of these backgrounds when they do cross the field.
3. Decoy discrimination, in which any flares are dropped by the target will be less effective because of the use of the small-field second-order servo.

Practical difficulties have generally discouraged switching field size in existing systems.

18.7.1.3. *Sensitivity Analysis.* A similar sensitivity analysis to that devised for a search system may be applied to a track system, that is, limiting optimized sensitivity defined by the system parameters of field of view, frame time (or servo response time), and number of detectors (as well by the physical parameters of optics diameter and detector's D^*). For purposes of this analysis, it will be assumed that the system is operating properly if quasistationary target tracking is maintained within the tracker's field of view under the influence of internal random noise and drift disturbances.

The track loop may be divided into three basic elements: an error sensor which generates an electrical signal as a function of target error angle; a demodulator which processes the signal into the form necessary to drive the gimbals, or its equivalent; and the gimbal drive, which generally acts as the smoothing element in the loop. The loop is closed from the gimbal drive to the error sensor.

Three primary classes of error-modulation techniques have been used: (1) amplitude modulation, where off-axis error is a function of percentage modulation and phase defines polar direction; (2) frequency modulation, where error is proportional to deviation and polar sense to phase of the deviation; and (3) pulse position, which is a track-while-scan technique.

It should be emphasized that, in a modulation error-sensing track system, the error information is contained in the modulation, not in the carrier. The modulation signal-to-noise ratio for maximum error signal as seen at the servo filter (which it is assumed is maintained at a constant closed-loop frequency by agc or clipping as required) that

determines whether tracking will be maintained. The signal-to-noise ratio required in the servo bandwidth may be defined by recognizing that a noise pulse in this bandwidth equal to the full modulation error signal will, to a first order, result in motion of the system to remove the target from the field of view. If it is desired, then, that in 100 time constants the probability for loss of track is to be 1%, this requires a probability of 10^{-4} per time constant. To find the signal-to-noise ratio corresponding to this probability, reference is made to the search false alarm analysis, modified to correspond to a carrier system. The number of positive peaks per integration time N, is found as follows:

$$T = 1/2\pi\,\Delta f$$

$$N = 0.64\,\Delta f$$

$$NT = 1/2\pi(0.64) = 0.25$$

Since either "positive" or "negative" peaks can cause loss of target, $2NT$ is used, or 0.5 peaks per servo time constant. The probability of a noise pulse to be used is thus 2×10^{-4}, and the signal-to-noise ratio required in the servo loop bandwidth is approximately 4.5. This ratio may then be traced back through the demodulation block to define required error sensing signal to noise. If the demodulation process is linear (synchronous), the computation is simple. The selection of a modulation frequency and bandwidth is not generally significant, and performance is defined by the servo bandwidth. For the more usual nonlinear demodulation process (AM detection, FM discriminator detection, etc.) the signal-to-noise ratio of the carrier system required to achieve 4.5 in the servo bandwidth is generally greater than that for the linear system and system performance is not optimized. This poor sensitivity match between optimized and realized systems stems from servo-loop stability considerations, *e.g.*, the sampled data problem in a track-while-scan system. To reduce the servo phase lag from the sampling process to less than 20°, the sampling frequency must be 10 times the servo bandpass. For a 10° lag, a ratio of approximately 20 is required. In real systems, the lags from low-frequency shaping and high-frequency cuts as well as gain-stability variation are such that all controllable lags must be kept low.

A track-while-scan system is usually designed with a threshold applied to detector output, after which error is found by pulse time or position in a demodulator. With this now linear (threshold) detection process, the result of the required high scan rate (10 times servo response) is to reduce the signal-to-noise ratio by $\sqrt{10} = 3.3$ compared with the optimized system which provides new information at the servo response rate. A similar level of problem exists in the other nonlinear detection systems (including AM and FM discriminator).

18.7.1.4. *Field of View, Servo, and Accuracy Analysis.* It has already been indicated that acquisition field size is defined by search-system target-pointing accuracy and that the track field may usually be made quite small. The factors defining track-field size include:

(1) Maximum target crossing velocity component at minimum range, $1/m$, defined from tactical analyses and provided as inputs to the infrared system design.

(2) Target maneuver capability, provided as inputs to the infrared system design.

(3) Imperfect stabilization under own vehicle motion resulting from infrared system stabilization dynamics and vehicle input motion. Vehicle motions are provided as inputs to the infrared system design.

(4) Biases, nonlinear friction, and other design limitations of the infrared system.

The track loop frequency response will generally be 6-12-6-12 although 6-18-6-12 and others have also been used. The first breakpoint from 6 db/octave to 12 db/octave occurs at a frequency low enough to exceed problem time ($1/\omega_1 >$ problem time). The second breakpoint (12 to 6) occurs at approximately one-third crossover (0 db gain) frequency and the 6 db slope exists for stability considerations. The ratio of $\omega_3/\omega_2 = 10$ and the last slope exists to eliminate high-frequency noise from the system. The ratio $\omega_3/\omega_2 = 10$ provides for system stability over normally expected system gain variations.

The problem of acquisition-error curve shape and servo-loop shaping to place the target in the track field of view will be considered under assumptions that:

(1) Acquisition and track modes have a common front end.

(2) Acquisition servo shaping is optimized separately from track shaping.

The common-front-end consideration dictates the desirability (in AM or FM systems) of full error modulation over the smaller track field, since this optimizes track accuracy and background discrimination. Thus, the desired error-curve shape in such systems would be linear as far as the track field, with full modulation at maximum expected track error, and then flat as far as the edge of the acquisition field. Background discrimination in track is aided by the nonlinear error-curve shape (compared with a linear error curve) since elements outside the track field contribute no more modulation than those within it.

In the pulse-position system, the common acquisition and track fields present no problem since a linear error curve may be indefinitely extended for acquisition, with no loss in track accuracy. Track-field sensitivity does, however, decrease with such an increased field. During track, background discrimination in the pulse-position system may be achieved by logical rejection of any error signal outside the track field.

If acquisition servo-loop shaping is to be optimized independently of track shaping, the acquisition closed-loop frequency is chosen by trading off sensitivity and acquisition time, utilizing maximum tolerable acquisition time to maximize sensitivity. The effect (on acquisition time) of the saturated error curve in AM and FM modulation systems is to increase acquisition time for a fixed closed-loop design frequency (assuming this design is based on the linear portion of the error curve). In addition, if the system must remain in the acquisition mode when it is within the track field, servo stability may be adversely affected unless a simple first-order design is used, since a 6-12-6-12 loop design will cause excessive overshoot at the reduced gain resulting from the saturated error curve. If logical switching to track is accomplished, the tendency to overshoot may, however, be successfully used to shorten acquisition time, provided proper damping of initial angular rate occurs when the switch to track is achieved.

18.7.2. Sample System.

(*A*) *System Input and Limitations.* The parameters around which the sample system are to be defined are:

(1) Target Radiation. 1000 effective watts per steradian in the detector wavelength band (3 to 5 μ).

(2) Target Motion. Maximum requirements of servo response will be based on a target at 1 n mi range with a crossing velocity of 1000 fps and a 4-*g* target turn at 1 n mi.

(3) Acquisition

(*a*) Field of view: 4° diameter

(*b*) Time: 0.3 sec

(4) Track Accuracy. 2 mrad under worst conditions

(5) Space. Allowable front-end cross section: 10 in. diameter

(*B*) *Front-End Description*

(1) Optics. Allowing for covering dome, clearance, and gimbals, an optics diameter of 7.5 in. may reasonably be achieved. An $f/1$ refractive or folded catadroptric system is required for short length to allow for maximum gimbal angle coverage. For background discrimination, optical resolution of 1 mrad or less to 2° off-axis is required. A bandpass filter in the general range of 3 to 5 μ is required. Expected net optical transmission over the band is about 35%.

(2) Detector. A cruciform detector array canted 45° to the horizon is used, subtending 4.5° along each axis by 1 mrad of detector width. The 45° angular tilt results in a scan which will not sharply cross the horizon, the primary background difficulty. Detector material is indium antimonide with a peak D^* in the band of 5×10^{10} cm $\text{cps}^{1/2}$ w^{-1}. To avoid excessive system complexity, the same detector and scanning should be used in acquisition and in track mode.

(3) Scanning. Image nutation over a 4° field is caused by rotation of an optical wedge or a tilted mirror.

(4) Track-Loop Servo Analysis. Servo requirements of the track loop are defined, and the adequacy of acquisition time is evaluated. To be consistent with a worst condition of 2-mrad error, linear dynamic lag is maintained at 1 mrad. The maximum angular rate may be calculated as 9.6° sec^{-1}, and maximum angular acceleration is approximately 1° sec^{-2}. Angular acceleration from a 4-g target turn at 1 n mi is 1.2° sec^{-2}. The effective velocity constant must then be 170 sec^{-1} at the "time" of maximum angular rate, and the acceleration constant must be 21 sec^{-2}. Frequency response of the 6-12-6-12 loop to achieve these K_v and K_0 values will result in an 8 rad sec^{-1} crossover frequency. The simplest acquisition computation to make is for a first-order system. The 6-12-6-12 system described here will have a faster response. For a target 2° off-axis at acquisition and an 8 rad/sec first-order loop, the time to reduce error to 2 mrad, or 0.059 of the initial value is approximately 0.3 sec; or to state this another way, specification reduction of acquisition error to less than 1 mrad requires approximately 0.4 sec.

With the 8 rad/sec cutoff frequency, the required information rate to achieve less than 10° phase lag at the servo closed-loop frequency corresponds to an angular frequency of about $(20)(8) = 160$ rad/sec, or $160/2\pi = 25$ pps. Using a pulse-position technique, with a small track field, two independent error pulses per scan are obtained. Thus, a scan rate of about 13 cps should be used. This scan rate is a significant parameter in defining the sensitivity of a pulse-position error-generation system.

(5) Sensitivity and Range Calculation. Having defined the significant front-end parameters, it is now possible to compute system sensitivity and range performance on the specified target. The expression for noise equivalent intensity (NEI) is:

$$\text{NEI} = \frac{\text{NEP}_c}{(A_o)(\epsilon_o)r}$$

where NEP_c is the noise equivalent power on the detector, A_o is the optical collecting area, ϵ_o is net optical efficiency, and r is the system response to the input signal.

$$\text{NEP}_c = \sqrt{A_c\,\Delta f}/D^*$$

where A_c is cell area and Δf is bandwidth, from which we can calculate

$$\text{NEI} = \frac{\sqrt{A_c\,\Delta f}}{D^*\,(A_o)(\epsilon_o)r}$$

Optimum bandwidth in the single-pulse system, assuming a sinusoidal pulse, is

$$\Delta f = \frac{1.84}{(2\pi)T} = \frac{0.3}{T}$$

With a 4° circle scanned at 13 cps and a 1-mrad detector width,

$$T = \frac{(0.001)(57.3)}{(\pi)(4)(13)} = 0.00035 \text{ sec}$$

Thus,

$$\Delta f = \frac{0.3}{0.00035} = 860 \text{ cps}$$

Low Pass: Cutoff in the present example is 100 to 200 cps. Response factor is derivable as:

$$r = 0.67$$

Cell Area: Using a single cell per arm,

$$A_c = \Omega f^2$$

where Ω = solid angle and f = focal length. If

$$\Omega = (0.001)(4.5/5.3) = 7.85 \times 10^{-5} \text{ sr}$$

and the focal length is 7.5 in.,

$$A_c = (7.85 \times 10^{-5})(7.5)^2(6.45) = 0.00285 \text{ cm}^2$$

Aperture area has been given as

$$A_{aperture} = \frac{\pi}{4}(7.5)^2(6.45) = 285 \text{ cm}^2$$

Optical efficiency has been given as:

$$\epsilon_o = 0.35$$

and

$$D^* = 5 \times 10^{10}$$

Thus,

$$\text{NEI} = \frac{\sqrt{(0.0285)(860)}}{(5 \times 10^{10})(285)(0.35)(0.67)} = 1.5 \times 10^{-12} \text{ w/cm}^2$$

Signal-to-Noise Ratio for Acquisition: A tolerable false alarm rate of 1 in 10 servo response times (once in 10/8 = 1.25 sec is assumed. Since acquisition does not require a high "hit" percentage of signal, 50% probability of detection is assumed.

With the 860-cps bandwidth, the number of signal peaks in 1.25 sec is:

$$N_{1.25} = (1.25)(0.775)(860) = 800$$

False alarm probability is then 1/800, or 0.0013. A threshold is then required at 4 times the rms noise. The signal-to-noise ratio for 50% probability of detection is also 4. Thus, a target which is to be acquired has a radiation at the system aperture of:

$$I = 4 \text{ NEI}$$

$$= 6 \times 10^{-12} \text{ w/cm}^2$$

Detectable Range Target: Assuming 30% atmospheric transmission, a 1-kw/sr target may be detected at:

$$r = \sqrt{\frac{J}{I}\,\tau_a}$$

$$= \sqrt{\frac{(1000)(0.3)}{6 \times 10^{-12}}}$$

$$= 7.1 \times 10^6 \text{ cm}$$

$$= 38 \text{ n mi}$$

Comparison of Sensitivity with a Servo Bandwidth-Limited System: The field in a servo bandwidth-limited system is effectively scanned once per servo response time, but a signal-to-noise ratio of about 6 is required. The system described here is thus compared with the servo bandwidth-limited system by the ratio:

$$\frac{\text{sensitivity-instrumented system}}{\text{servo bandwidth-limited system}} = \frac{6}{4}\sqrt{\frac{8}{2\pi(13)}} = 0.48$$

A servo bandwidth-limited system would thus have twice the sensitivity of the instrumented system.

(*C*) *Block Diagram.* The block diagram of the sample system is shown in Fig. 18-5. By tracing through the significant elements of this diagram, the trade-off logic is demonstrated.

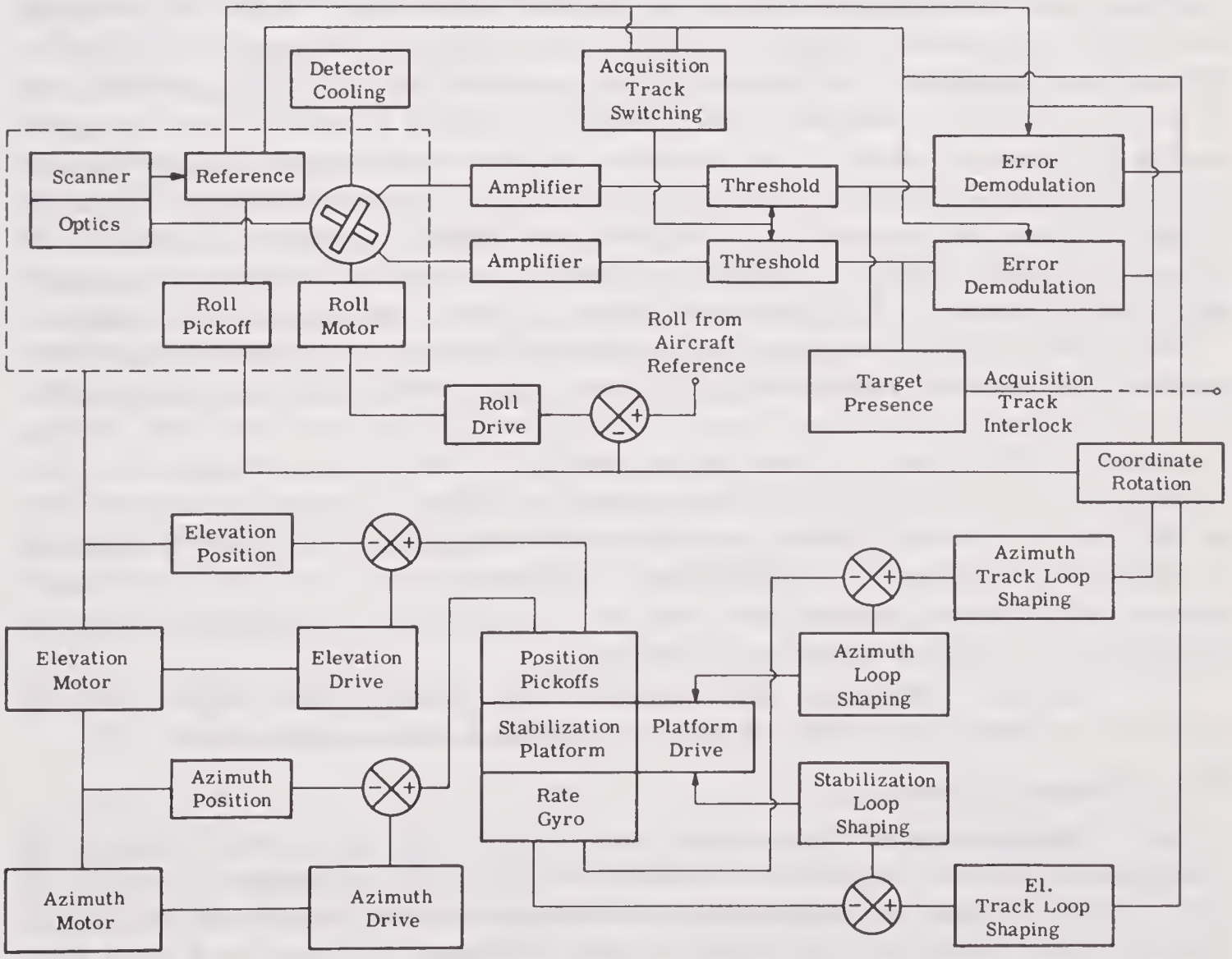


FIG. 18-5. System block diagram for track.

(1) Error Generation. Cruciform-detector outputs will be separately amplified in gain-stabilized amplifiers with strong signal limiting to avoid amplifier blocking by saturation. Amplifier outputs drive threshold circuits to remove noise from the signal at a level defined by a tolerable false alarm rate. The threshold outputs drive error demodulators which define target error position. These error demodulators derive scan-position information from a reference generator located at the nutation drive in the optical system. System logic provides (in the acquisition track switching) that, after initial acquisition (as defined by error dropping below a critical threshold), the threshold circuit will not pass signals corresponding to errors greater than those expected in track (say, 1/4 degree).

(2) Target Presence Generation. When enable logic initiates acquisition, an interlock based on detection of a target in both detector channels is provided. When the switch to small-field track mode is initiated, the target-presence circuitry operates only on track field signals. Loss of signal for any reason should probably be used to initiate a return to the search mode.

(3) Stabilization of Track. Two stabilization components are used in this sample system. The less conventional one is roll stabilization of detectors for background discrimination to avoid severe horizon crossing, which may occur in maneuvers even with the 45° rotated detectors. This roll stabilization need be only approximate. It receives its command from aircraft instruments. The effect of detector roll position must be considered in error generation (reference generator affected) and in coordinate rotation of the error signals necessary to drive the conventional azimuth-elevation gimbals.

The more conventional stabilization is that about the line of sight to the target to maintain track with respect to the tracking aircraft's maneuvers. Before one considers the analytical parameters of this stabilization, the problem of placement of stabilization gyros should be settled. Conventionally, they are placed on the sighting platform, and if space and heat-sink problems permit, every effort should be made to locate the gyros on the IR telescope. If this is not possible, a separate stabilization platform may be provided to which the sighting head is slaved with an aided rate command. This technique has been used successfully on at least one infrared equipment. Sight line remains to the target since the position loop provides at least one integration to remove any slaving errors. In Fig. 18-5, stabilization is shown using an auxiliary platform.

Analytical treatment is independent of the gyro location. An electric-motor-driven geared gimbal system is adequate for this class of equipment. With conventional gear trains, stabilization closed-loop frequency is generally limited to 8 to 15 cps. In this treatment, a 65 rad/sec (~10 cps) loop is assumed. The basic loop response is shown in Fig. 18-6. The "long-period" velocity constant is 980 sec^{-1}, and acceleration constant is 1950 sec^{-2}. The stabilization loop need not be a high-order loop below frequencies at which the track loop is tightly closed. The acceleration constant of 1950 sec^{-2} means that a 1-mrad dynamic error may be maintained under acceleration inputs of approximately 110°/sec^2 as seen at the gimbals.

(4) Cooling. Depending upon mission time, logistics, and weight trade-off penalities, an open-loop or closed-loop cooling system decision may be made.

18.8. Mapping Systems

18.8.1. Requirements. This section does not deal with the problem of extracting the information from a map (often the critical element in mapping-system utility) but only with maximizing map-information content, which is achieved by use of a small resolution-area detection of minimum contrast combined with scan of a large field. As usual, reconciling these quantities involves compromise.

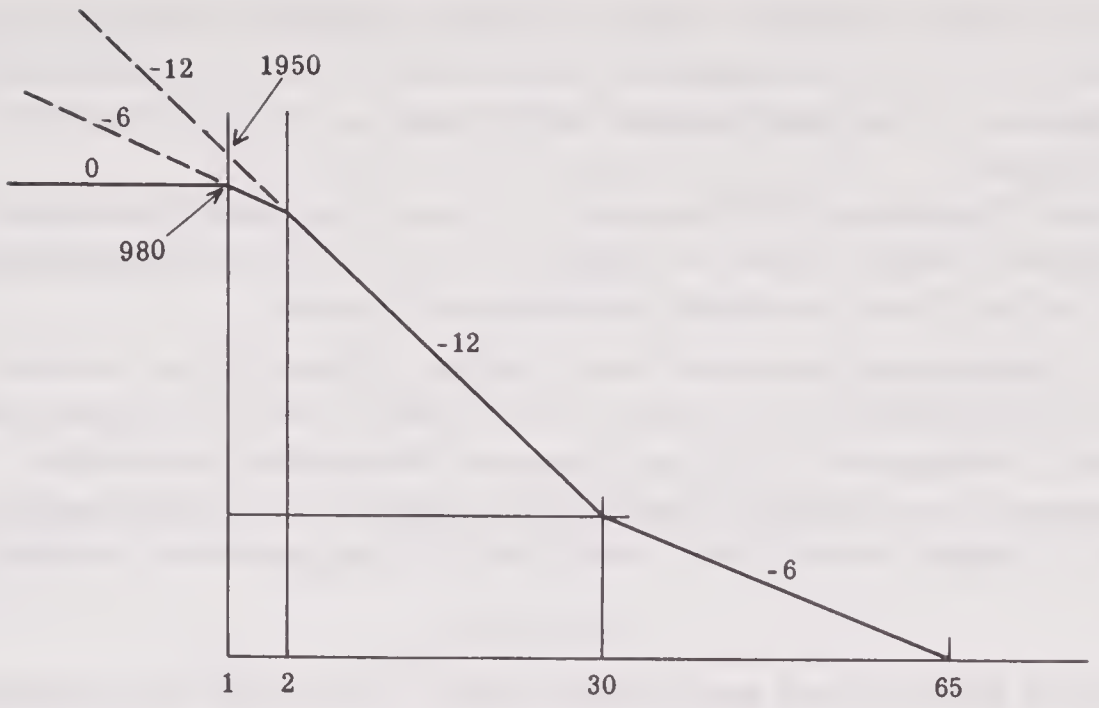


FIG. 18-6. Basic loop response.

18.8.2. Basic Sensitivity Equation.

18.8.2.1. *Detector-Noise-Limited System.* The derivation of the sensitivity equation for a detector-noise-limited system may always be started from the simple detector-sensitivity expression:

$$\mathrm{NEP}_c = \sqrt{A_c\,\Delta_f}/D^*$$

and the resultant overall system expression:

$$\mathrm{NEI} = \frac{\mathrm{NEP}_c}{(A_o)(\epsilon_o)r} = \frac{\sqrt{A_c\,\Delta f}}{D^* A_o \epsilon_o r} = \frac{FD}{D^*(\pi D^2/4)\epsilon_o r}\sqrt{\frac{0.65\Omega}{T_f}} = 1.1\,\frac{F}{D^* D \epsilon_o r}\sqrt{\frac{\Omega}{T_f}}$$

The detectable signal may be found by computing the incremental signal at the aperture and setting it equal to the NEI multiplied by some constant. This constant for detection in a single resolution element would have to be same as for a point detection system (about 4 to 7 times the NEI). However, with the element-to-element integration of a map, larger-area information is derivable at a lower signal-to-noise ratio down to below unity in a single element.

The incremental radiation dW may be found from:

$$W = \epsilon\sigma T^4$$

$$dW = \sigma T^4\,d\epsilon + 4\epsilon\sigma T^3\,dT$$

$$= W_{bb}\,d\epsilon + 4\epsilon W_{bb}\,\frac{dT}{T}$$

where ϵ = emissivity

σ = Stefan-Boltzmann constant

W_{bb} = blackbody radiant emittance ($\epsilon = 1$)

Normal system use is limited by detector response and atmospheric transmission to a spectral band less than that essentially encompassing total radiation. Also, atmospheric transmission may further attenuate the signal. The incremental radiation dH at the aperture in a resolution-element solid angle ω is then:

$$dH = \left(W_{\lambda_1\text{-}\lambda_2}\,d\epsilon + 4\epsilon W_{\lambda_1\text{-}\lambda_2}\,\frac{dT}{T}\right)\tau_A\,\frac{\omega}{\pi} \qquad \text{w cm}^{-2}$$

where $W_{\lambda_1\text{-}\lambda_2}$ is total blackbody radiation (over a hemisphere) per unit area in the interval from λ_1 to λ_2, τ_A is atmospheric transmission in the λ_1-λ_2 band, and ω/π is the ratio of radiation in the instantaneous field of view, ω, to the full radiation of the body over a hemisphere.

For a given background environment, it is seen that, within practical limitations such as detector response time, the product $\omega\, d\epsilon$ or $\omega\, dT$ permits a trade-off between resolution and minimum detectable temperature difference.

18.8.2.2. *Background-Radiation-Noise-Limited System.* Some infrared detectors are background-radiation-noise-limited when they view an ambient temperature background. A design for this type of detector is derived in this section.

It is assumed that the detector has cold shielding, allowing a solid angle of radiation acceptance, θ, which is twice that corresponding to the geometric f number, $\mathcal{F}$; that is

$$\theta = 2/\mathcal{F}^2$$

It will also be assumed that temperatures inside and outside the instrument are the same and that by reflection of similar temperature sources the net background within the angle ω may be approximated as a blackbody.

If the instantaneous and total fields of view of the system are ω and Ω and the frame time is T_f, detector area A_c is related to ω, $\mathcal{F}$, and D_o by:

$$A_c = \omega \mathcal{F}^2 D^2$$

The dwell time on target is:

$$t = \frac{\omega T_f}{\Omega}$$

Again using a bandwidth

$$\Delta f = \frac{0.65}{t} = \frac{0.65\Omega}{\omega T_f}$$

The rms noise due to radiation noise in a bandpass Δf is the same as that in an integration time

$$T_i = \frac{1}{2\Delta f} = \frac{\omega T_f}{1.3\Omega}$$

The background radiation in the sensitive band λ_1 to λ_2 received by the detector through a solid angle θ is

$$W = W_{\lambda_1\text{-}\lambda_2} A_c \frac{\theta}{\pi} = W_{\lambda_1\text{-}\lambda_2} \omega \mathcal{F}^2 D^2 \left(\frac{2}{\pi \mathcal{F}^2}\right)$$

$$= \frac{2}{\pi} W_{\lambda_1\text{-}\lambda_2} \omega D^2$$

If the detector has a quantum efficiency Q, the number of effective background photons received in the integration time, T_i, is:

$$N_B = Q \frac{W}{hc/\lambda_c} T_i$$

$$= \frac{Q \frac{2}{\pi} W_{\lambda_1\text{-}\lambda_2} \omega D^2}{hc/\lambda_c} \frac{\omega T_f}{1.3\Omega}$$

$$= 0.49 \frac{Q \lambda_c W_{\lambda_1\text{-}\lambda_2} \omega^2 D^2 T_f}{hc\Omega}$$

where λ_c = center wavelength of λ_1-λ_2 and hc = the number of noise photons in the integration time, T_i, is:

$$N_N = \sqrt{N_B} = \omega D \sqrt{(0.49)\frac{Q\lambda_c W_{\lambda_1\text{-}\lambda_2} T_f}{hc\Omega}}$$

The detector NEP is:

$$\text{NEP}_c = \frac{N_{N_{T_i}}}{T_i}\frac{hc}{Q\lambda_c}$$

$$= \frac{1.3\Omega\omega D}{Q\omega T_f}\sqrt{0.49\frac{Q\lambda_c W_{\lambda_1\text{-}\lambda_2} T_f}{hc\Omega}}\frac{hc}{\lambda_c}$$

$$\text{NEP}_c = 0.91D\sqrt{\frac{hc}{\lambda_c}\frac{\Omega}{QT_f}W_{\lambda_1\text{-}\lambda_2}}$$

Then:

$$\text{NEI} = \frac{\text{NEP}_c}{A_o\epsilon_o r} = \frac{\text{NEP}_c}{\frac{\pi D^2}{4}\epsilon_o r}$$

$$= \frac{0.91D\sqrt{\frac{hc}{\lambda_c}\frac{\Omega}{QT_f}W_{\lambda_1\text{-}\lambda_2}}}{\frac{\pi D^2}{4}\epsilon_o r}$$

with $r = 0.9$

$$= (1.3)\frac{\sqrt{\frac{hc}{\lambda_c}\frac{\Omega}{QT_f}W_{\lambda_1\text{-}\lambda_2}}}{D\epsilon_o}$$

18.8.2.3. *Multiple Detector System.* If a multiplicity of detectors is used, the total field searched per frame time is divided by N_D, the number of detectors, and computed sensitivity is improved by $\sqrt{N_D}$, or

$$\text{NEI} \propto \frac{1}{\sqrt{N_D}}$$

It must be noted that if multiple detectors are used in a mapping system (and they may be desirable for increasing resolution and decreasing the required detector response time), and if a striped map is to be avoided, detector responsivities as seen at the output of the signal processing circuitry must be matched.

18.8.2.4. *Edge Resolution vs. Signal Strength.* The term instantaneous resolution is related to the minimum detectable signal level and is limited to the resolution area of the detector. In the presence of a strong signal, however, full response need not be achieved to define an edge. To a first order then, if data-reduction techniques are used to best advantage, the usable resolution solid angle varies inversely as the contrast signal-to-noise ratio so that a doubling of signal-to-noise ratio contains information allowing half the resolution of an isolated edge.

18.8.3. Background-Limited D^*-Q Trade-Off. To find the D^*-Q relationship defining the value of detector D^* that is equivalent to a specific value of Q in the system configuration defined in Section 18.8.2, the two NEI expressions are equated.

$$\frac{1.1\,F}{D^*D\epsilon_o}\sqrt{\frac{\Omega}{T_f}} = 1.3\,\frac{\sqrt{\dfrac{hc}{\lambda_c}\dfrac{\Omega}{QT_f}\,W_{\lambda_1\text{-}\lambda_2}}}{D\epsilon_o}$$

With $F = 1$, $W_{\lambda_1\text{-}\lambda_2} = (0.031)(0.4) = 0.012$, and $\lambda_c = 10\,\mu$ ($8-15$-μ band at 0°C),

$$\frac{D^*}{\sqrt{Q}} = 5.4 \times 10^{10}$$

18.8.4. Sample System.

18.8.4.1. *Specific Requirements.* The basic requirement may be stated simply as defining a low-altitude mapping system to operate at 1000 feet and at a velocity of 800 fps and to provide a thermal map, in the 8-15-μ atmospheric window, of 0.5 mrad resolution (6 in. at a nominal 1000-ft altitude). The desired signal detectivity in the extended area (unity signal-to-noise ratio) is 1°C against a nominal 0° C background of 0.5 emissivity. The system should be capable of operation up to 10,000 ft at 800 fps with constant optical resolution and the stabilization required for production of good map information. Stabilization and image motion compensation must provide for 10% image smearing line-to-line.

18.8.4.2. *Sensitivity Computation.* As the basis for a realizable system, a D^* corresponding to a quantum efficiency of 0.1 will be assumed. Neglect time constant for the moment; the optics diameter necessary to achieve a signal-to-noise ratio of unity for a 1°C change in temperature is found by equating the signal to the NEI.

$$\frac{1.1\mathcal{F}}{D^*D\epsilon_o}\sqrt{\frac{\Omega}{T}} = 4\epsilon W_{\lambda_1\text{-}\lambda_2}\frac{dT}{T}\,\tau_A\,\frac{\omega}{\pi}$$

The values for the parameters are:

$\mathcal{F} = 1$

$D^* = 1.7 \times 10^{10}$

$\epsilon_o = 0.3$

$\Omega = (0.5 \times 10^{-3})(2\pi)$, assuming for simplicity a continuous single element with ~50% dead time above the horizon scan

T is found from v/h and the detector dimension along the flight line:

$$T = \frac{0.5 \times 10^{-3}}{800/1000} = 0.625 \times 10^{-3}\ \text{sec}$$

$\epsilon = 0.5$

$W_{\lambda_1\text{-}\lambda_2}$ for 0°C and 8 to 15 $\mu = (0.031)(0.4) = 0.0124$ w cm^{-2}

$$\frac{dT}{T} = \frac{1}{273}$$

$\tau_A = 0.6$

$\omega = (0.5 \times 10^{-3})^2 = 0.25 \times 10^{-6}$

Solving for D,

$$D = 112\ \text{cm} = 44\ \text{in.}$$

This is certainly an unreasonable dimension.

18.8.4.3. *Use of Multiple-Detector Array.* The aperture is large; the dwell time is short. Thus,

$$t = \frac{\omega}{\Omega} T = \frac{0.25 \times 10^{-6}}{(2\pi)(0.5 \times 10^{-3})} (0.625 \times 10^{-3})$$

$$= 5 \times 10^{-8} \text{ sec}$$

From both a sensitivity and a time-constant standpoint, it is desirable to use a detector array.

If a 10-in. aperture is to be used, 20 detectors are required and dwell time on target is 10^{-6} sec.

18.8.4.4. *System Configuration.* (1) Front End. The optical system and detector array are mounted on a roll-pitch stabilized platform. Scan is by a 45° tilted mirror continuously rotating about an axis along the aircraft heading and at a speed defined by v/h input, discussed below. The nominal operating rotation speed will be 80 rps (800 fps at 1000 ft and 20×0.5 mrad = 10 mrad per scan). To keep the image properly placed with respect to the detector, a dove prism which rotates an image at twice its own angular rate is used.

Detector cooling may or may not be on the stable platform depending on the trade-off between liquid-transfer problems and platform size and weight. The detector array is packaged in a single dewar and aligned to sight along the flight path. The detectors may be staggered in the focal plane rather than in a single line to avoid the problem of spacing between detectors. A cold shield is used to isolate the spectral interval and define the field of view.

(2) Signal Processing. The main problem in the signal processing is that of maintaining detector-to-detector consistency. To do this, operation must be below detector cutoff frequency so that detector time constant is uniquely defined for all elements. Amplifier frequency responses must be closely matched using stabilized circuit techniques. The separate channel outputs may either be separately recorded by multichannel means or may be commutated by high-speed gates sampling 3 times per dwell time. Because of the resolution degradation as scan departs from nadir view a programmed reduction in the number and choice of detectors for recording may be used.

(3) Stabilization Requirements. Long-time stabilization must be better than 2 mrad and short-term scan-to-scan stabilization about 0.1 of the system resolution, or 0.05 mrad. This means 0.05 mrad in 1/80 sec, or 4 mrad sec^{-1} rate error. At 10,000 ft altitude, 0.4 mrad sec^{-1} is required. To achieve good short-period stabilization, a gearless torqued gimbal structure with self-contained gyros will be used. Long-period vertical reference information may be transmitted from aircraft inertial reference equipment.

(4) Image Motion Compensation

(*a*) *Requirement.* IMC in a high-speed scanning system is a less critical problem than in a photographic system. With a single detector, 10% scan-to-scan resolution smearing corresponds to 10% accuracy in v/h. With 20 detectors, smearing of 10% on one detector image means 0.5% v/h accuracy. Drift-angle accuracy follows the same pattern. With a single detector, 10% drift corresponds to an angle of approximately 6°. With 20 detectors, the requirement is 0.3° to achieve 10% scan to scan misalignment.

(*b*) *Implementation.* v/h compensation will be made a part of the system by varying scan rate to achieve proper matching of line to line scanning. Drift compensation will be made a part of the system by skewing data output.

(*c*) v/h *and Drift Measurement*

1. *External Source:* If the aircraft has a radar altimeter and doppler, inertial, or doppler-inertial navigation, v/h and drift angle to the required accuracy may be derived from these sources.

2. *Passive* v/h *and Drift*: The mapping system itself, with some additional focal-plane detectors and additional processing, offers a self-contained source of v/h and drift information. v/h is computed by correlating data sampled from the central angle of the scan (directly below the aircraft) for overlap or underlap on two successive scans, using v/h "overlap" detectors added for this purpose. Drift angle is computed by correlation techniques looking for scan-to-scan skewness in data sampled from directly below the aircraft on detectors separated by a large angle along the flight path and caused to overlap by v/h to cover the same area on successive scans. The v/h and drift error signals may thus be generated to vary scan speed and drift angle.

Chapter 19

INFRARED MEASURING INSTRUMENTS*

CONTENTS

19.1. Radiometers 758
19.1.1. Reference Radiation Level 759
19.1.2. Commercially Available Radiometers 761
19.1.3. Normalization Methods 761
19.2. Spectroradiometers 763
19.3. Monochromators 764
19.4. Spectrometers and Spectrophotometers 766
19.4.1. Infrared Grating Spectrometry 766
19.4.2. Commercially Available Spectrometers and Spectrophotometers 767
19.5. Interferometers 776
19.5.1. Rayleigh Interferometer 776
19.5.2. Michelson Interferometer. 777
19.5.3. Twyman-Green Interferometer 777
19.5.4. Fabry-Perot Interferometer 778
19.5.5. Spherical Fabry-Perot Interferometer 779
19.5.6. Lummer-Gehrcke Plate 779
19.5.7. Spectral Transmittance of Interferometers 780

*Material prepared by the technical writing staff of McGraw-Hill, Inc.

19. Infrared Measuring Instruments

19.1. Radiometers

A radiometer is a radiation-measuring instrument having substantially equal response to a relatively wide band of wavelengths in the infrared region. Radiometers measure the difference between the source radiation incident on the radiometer detector and a radiant energy reference level.

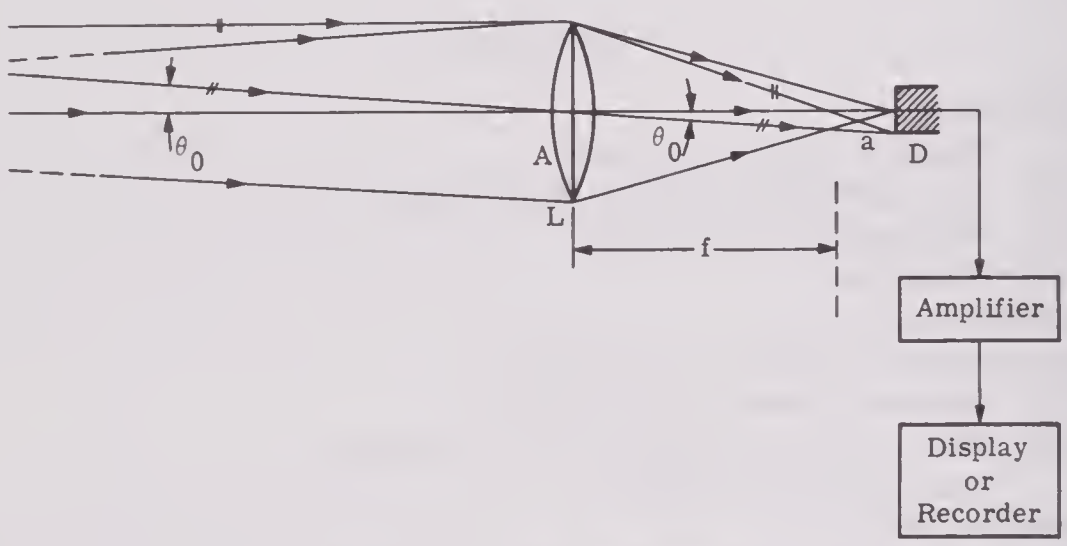


FIG. 19-1. Basic radiometer.

The basic design of a simple radiometer is shown in Fig. 19-1, where

L = collecting optics, which forms the circular aperture stop of area A

D = dectector element, which forms a circular field stop of area a

θ_0 = half-angle, measured in radians ($\Omega = 2\pi\ (1 - \cos\theta_0) \approx \pi\ \theta_0^2 \approx a/f^2$); Ω is the solid angle, is steradians, of the corresponding conical field

f = focal length of radiometer

All radiometers and radiometric measuring instruments contain at least the following three essential components:

(*a*) A detector element, which converts changes in incident electromagnetic radiation into variations of an easily measured property, usually an electrical signal.

(*b*) An optical system, which determines the combination of receiving aperture and angular field of view through which radiation is collected, thus delineating the amount of radiation to which the radiometer responds. The optical system includes the sensitive surface of the detector.

(*c*) An amplifier and output indicator, usually electronic, to transform the output of the detector element into the desired form of presentation.

19.1.1. Reference Radiation Level. Both absolute and relative infrared radiation levels may be obtained with a radiometer. Since absolute levels of radiation are defined with respect to absolute zero, the level of radiant power incident upon the detector must be compared with a known reference level to derive the absolute radiant power level.

The ultimate accuracy of a radiometer is determined by the form of reference radiation level used. The detector itself (dc radiometer) or a radiation chopping system (ac radiometer) can be used to provide the reference radiation level. For absolute measurements, the chopping system is preferred.

Radiometers that measure the difference in radiation from two neighboring spatial positions provide relative information only as no reference level exists. Radiometers used to compare any element of a large area to the average radiation associated with the entire area can supply an absolute measurement, providing the average intensity is known and used as a reference level. Similarly, in the time domain, the power level due to radiation at one instant may be compared to that of a previous instant or to an average associated with all past measurements.

The characteristics of three different type radiometers, including the form of reference level they employ, are listed in Table 19-1.

TABLE 19-1. PRINCIPAL CHARACTERISTICS OF DIFFERENT TYPE RADIOMETERS [5]

Type	*Detector*	*Detector Response Speed (Time Constant)*	*System Speed*	*Reference Radiation Level*	*Radiation Measurements*	*Electrical Signal*
DC	Thermopile	2 sec	2 sec	Emissivity and temperature of thermopile	Difference of source radiation and thermopile radiation	Electromotive force from compensated dc thermopile
AC Blackened Chopper	Thermistor	1 msec	25 msec	Emissivity and temperature of blackened chopper	Difference of source radiation and blackened chopper radiation	AC signal from compensated thermistor bolometer bridge
AC Chopper Mirror	Thermistor	1 msec	Adjustable, 16 msec to 1.6 sec	Temperature of blackbody reference (emissivity = 1.0); temperature controlled or monitored within 0.2°	Difference of source radiation and reference blackbody radiation; null method may be used with temperature-controlled blackbody	AC signal from compensated thermistor bolometer bridge; null detection method when using controlled blackbody

19.1.1.1. *Detector Energy Level Used as Reference Radiation (DC Radiometer).* A radiometer that uses the detector as a radiation reference level is usually referred to as dc radiometer. This terminology is applicable because the instrument measures a change in the dc electrical properties of a thermoelectric or bolometric infrared detector.

DC radiometers are subject to drift because the reference level is determined by the temperature of the detector. Drift can often be tolerated when the temperature of the target is much higher than the detector's ambient temperature. Significant errors can result, however, when the target temperature approaches ambient temperature. The

typical response speed (time constant) for instruments using thermopile detectors is approximately 2—4 sec where the response speed is time required for the instrument to reach *1/e* of the final response.

19.1.1.2. *Chopper Used as Reference Radiation Level (AC Radiometer).* Chopper or ac radiometers utilize an ac output from the detector for signal processing. They are particularly suitable for absolute radiation measurement and do not have the drift problems associated with dc radiometers. The electrical output of this type radiometer is proportional to the difference between radiation falling upon the detector from the source within its field of view and that of a blackened chopper blade or a controlled reference blackbody. In the latter case, a chopper mirror alternately directs radiation from the source and the reference blackbody onto the detector.

Blackened Chopper. Figure 19-2 shows a blackened chopper where the detector alternately sees the source image and the blackened chopper. The temperature and emissivity of the blackened chopper determine the reference radiation level. At wavelengths out to about 1 μ, a very stable reference level is obtained by ensuring that the chopper has a uniformly coated black surface. Even out to about 3 μ, chopper temperature is not usually a critical factor because at ambient temperatures there is little radiation in this region from a blackbody or a graybody. At wavelengths longer than about 3 μ, however, the effects of variations in chopper temperature and emissivity become a serious consideration.

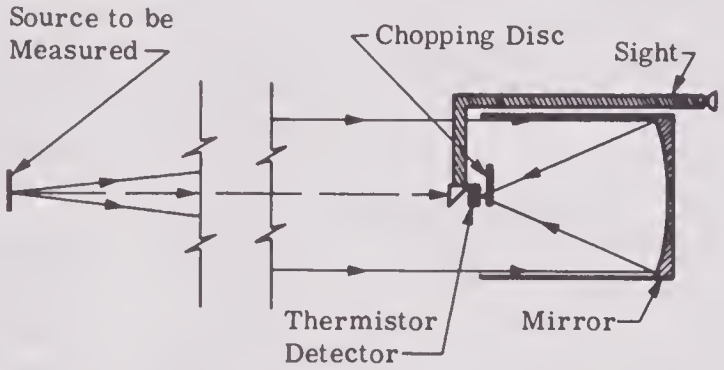


FIG. 19-2. Blackened chopper schematic diagram.

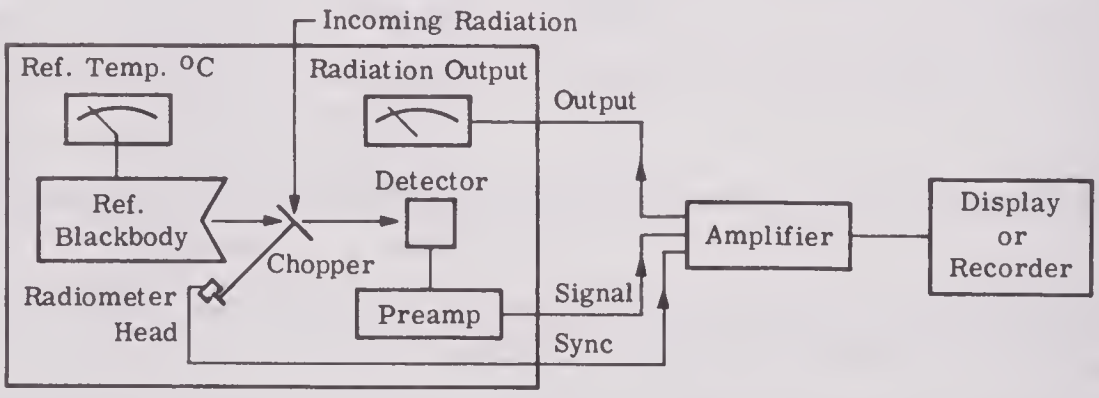


FIG. 19-3. Chopper mirror block diagram.

Chopper Mirror. Emissivity effects can be minimized by using a highly polished chopper blade. However, incident radiation reflected from the surrounding area to the detector by the polished chopper must be controlled. This is done in the radiometer shown in Fig. 19-3, where reflected incident radiation is controlled by a reference blackbody and a chopper mirror.

A small reference blackbody is placed in such a position that, by specular reflection from the chopper mirror, radiation from the reference blackbody and the source image are alternately directed onto the detector. This not only provides a stable source of reference radiation but also an adjustable one when temperature controls are provided for changing the reference blackbody temperature. In addition, if the reference blackbody source is adjustable, its radiance can be matched to the source radiance and the detector used as a sensitive null or quantitative error detector.

19.1.2. Commercially Available Radiometers. Table 19-2 lists some of the characteristics of most commercially available radiometers.

19.1.3. Normalization Methods. By careful design, some instruments achieve quite uniform responsivity over wide but finite spectral regions. In many cases, however, responsivity is not uniform even within the band, due to the combined effects of the spectral characteristics of the components (mirrors, lenses, filters, prisms, transducers, etc). Also, the source spectral distribution is almost never constant. In general, the output signal produced by an incident beam of radiant power

$$P = \int P_\lambda \, d\lambda \qquad \text{w} \tag{19-1}$$

is given by

$$V = \int P_\lambda R_p \, d\lambda \qquad \text{v} \tag{19-2}$$

where $R_p = R_p\,(\lambda)$ is the spectral power responsivity in volts per watt. Even where $R_p\,(\lambda)$ is completely known from calibration measurements, it is not possible to solve Eq. (19-2) uniquely for $P_\lambda\,(\lambda)$. Consequently, it is also not possible to determine $P = \int p_\lambda \, d\lambda$. However, if the *relative* spectral distribution $p_\lambda\,(\lambda)$ in the incident beam is known, the measurement will establish the scale factor P_s, which satisfies

$$P_\lambda\,(\lambda) = P_s p_\lambda\,(\lambda) \qquad \text{w} \tag{19-3}$$

from the relation

$$P_s = V \Big/ \int p_\lambda \, R_p \, d\lambda \qquad \text{w} \tag{19-4}$$

On the other hand, without any information about the spectral distribution of the incident beam, the output of the instrument can be compared meaningfully only with measurements made with instruments having exactly the same relative spectral responsivity characteristic.

Fortunately, it usually happens that there is at least some information about the spectral distribution $p_\lambda(\lambda)$, although it may be only approximate and often is only implicit and not clearly recognized. For example, from the known physical characteristics of a target—its material, its approximate temperature, etc—it is often possible to judge the general shape of the curve of $p_\lambda(\lambda)$ and hence to be able to estimate its relation to the spectral curves of other targets in order to interpret comparisons between measurements of two such targets by spectrally selective radiometers. In order to eliminate differences in other instrumental characteristics (which also affect responsivities) for the purpose of making such comparisons, normalization methods are often employed in reducing the data [1,2].

Note that, although this discussion is all in terms of radiant power P and power responsivity R_p, all of it applies equally well to the other radiometric quantities for an incident beam, H and N, and the corresponding responsivities, and to the source quantities J and W, derived from measured values of incident P, H, or N, as well. The latter, however, involve additional uncertainty because of the problem of evaluating the attenuation between source and instrument.

TABLE 19-2. CHARACTERISTICS OF COMMERCIALLY AVAILABLE RADIOMETERS

Name	*Spectral Range* (μ)	*Temperature Range* (°K)	*Detector Type*	*Optics*	*Number of Filter Positions*	*Max. Field of View* (*Mils*)	*Range of Focus* (*cm*)	*Focal Length* (*cm*)	*Effective Focal Ratio* (*f/no*)	*Resolution* (*Mils*)	*Accuracy* (°K)
Barnes*											
R-4F1, R-4F2	0.3-1000	–	TB§	Cassegrain	8	12.0 × 12.0	10-∞	20	2.5	0.1	–
R-8B1, R-8B2	0.3-1000	263-363	TB	Cassegrain	6	8.0 × 8.0	10-∞	30	1.9	0.1	–
R-8D1, R-8D2	0.3-1000	263-363	–	Cassegrain	6	8.0 × 8.0	10-∞	30	1.9	0.1	–
R-12A1	0.3-1000	–	TB	Pfund	–	8 × 8	250-∞	56	1.2	0.2	2
R-4D1	0.3-1000	Ambient-3873	TB	Cassegrain	–	2.5 × 2.5	5-∞	20	2.5	–	–
R-4K1	–	–	TB	Cassegrain	–	51 × 51	25-∞	30	f/3	–	–
Block†	–	–	–	–	–	–	–	–	–	–	–
E-2D	–	–	Pbs	C_aF_z	5	85 × 85	–	–	–	–	–
E-5	0.5 to 40		TB	KRS-5	5	85 × 340	–	–	–	–	–
Williamson‡	–	–	–	–	–	–	–	–	–	–	–
Stoll-Hardy HL	–	Ambient to ±373	TB	–	–	20°	–	–	–	–	–
Stoll-Hardy HL Industrial Model	–	255 to 487	TB	–	–	14°	–	–	–	–	–

*Barnes Engineering Co.
†Block Associates, Inc.
‡Williamson Development Co.
§Thermistor bolometer

By choosing arbitrarily a normalizing constant or scale factor R_n, which satisfies $R(\lambda) = R_n r(\lambda)$, Eq. (19-2) can be written

$$V = R_n \int P_\lambda r \, d\lambda = R_n P_n \qquad \text{v} \tag{19-5}$$

where

$$P_n = V/R_n \qquad \text{w} \tag{19-6}$$

is called the normalized radiant power in the incident beam.

19.1.3.1. *Normalization to the Peak.* If the choice of normalization constant is

$$R_n = R(\lambda_{max}) \qquad \text{v w}^{-1} \tag{19-7}$$

where λ_{max} is the wavelength at which $R(\lambda)$ is a maximum, as with most narrowband systems, then

$$P_n = \int P_\lambda r \, d\lambda \qquad \text{w} \tag{19-8}$$

is called the peak-normalized power in the beam or the power normalized to the peak responsivity. From Eq. (19-6) and (19-7), it is apparent that this means that P_n watts of radiant power, if concentrated at the wavelength λ_{max}, would produce the same output V as the actual beam. Hence, peak-normalized power is sometimes called the "effective" [3] or "equivalent" power at λ_{max}.

Furthermore, if the peak-normalized spectral bandwidth of the instrument is arbitrarily defined as

$$\Delta\lambda_n = [1/R(\lambda_{max})] \int R(\lambda) \, d\lambda \qquad \mu \tag{19-9}$$

it may be used to compute

$$P_{\lambda n} = P_n/\Delta\lambda_n \qquad \text{w } \mu^{-1} \tag{19-10}$$

the peak-normalized spectral radiant power assigned to λ_{max}.

19.1.3.2. *Other Methods of Normalization.* There are situations where normalization to the peak is inappropriate because there is no wavelength λ_{max} for which there is a single predominating maximum value of $R(\lambda)$. This is more often so with broadband systems. It may then be more appropriate to normalize to the average, by choosing

$$R_n = \left[\int_{\lambda_1}^{\lambda_2} R(\lambda) \, d\lambda \right] \Big/ (\lambda_2 - \lambda_1) \qquad \text{v w}^{-1} \tag{19-11}$$

Then the normalized bandwidth

$$\Delta\lambda_n = (1/R_n) \int_{\lambda_1}^{\lambda_2} R(\lambda) \, d\lambda \qquad \mu \tag{19-12}$$

is equal to the entire wavelength interval of interest $(\lambda_2 - \lambda_1)$ over which the average is computed in Eq. (2-11). The foregoing, and still other normalization methods and their implications, are discussed in greater detail in [1] and [2]. Because of the variety of possible normalization methods, it is important to specify clearly the method used when normalized values for measurement results are reported.

19.2. Spectroradiometers

Spectroradiometers are used to obtain an absolute measurement of the spectral variation of a source radiometric quantity within a very narrow waveband. Dispersing

elements such as prisms, diffraction gratings, or other optical elements can be used to produce the spectra.

Radiometers which rapidly sequence through a set of narrow filters also can be called spectroradiometers. The filter is often sequenced by the rotation of a filter wheel. If the speed of rotation is sufficiently high, the dwell time for any one filter can approach the response time of the unfiltered radiometer. In such cases, the response time should be determined under dynamic rather than static conditions.

The basic design of a simple prism or grating spectroradiometer is shown in Fig. 19-4, where

L = collecting optics
S_1 = entrance slit of monochromator acting also as the field stop
C = collimating optics
Δ = dispersing element (prism or grating)
F = refocusing optics
S_2 = exit slit of monochromator
D = detector element

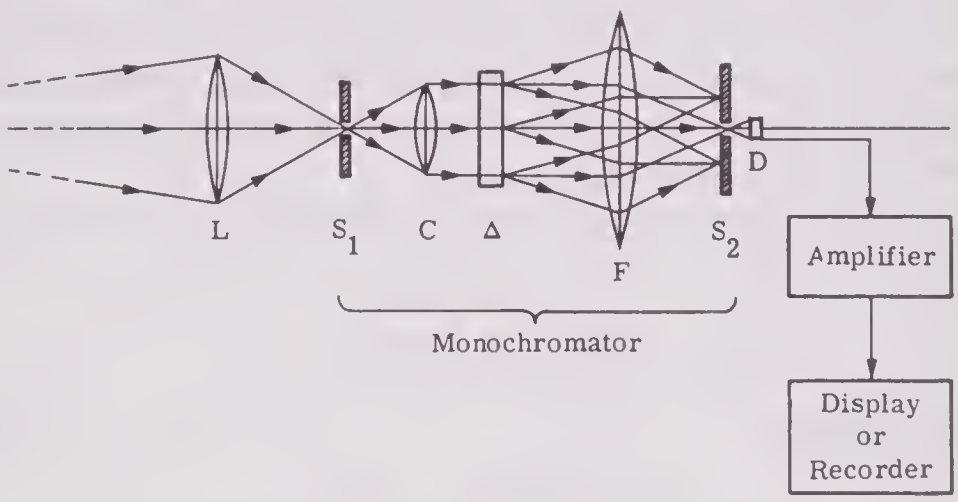


FIG. 19-4. Basic spectroradiometer.

The essential components of any prism or grating spectroradiometer are the same as those of any radiometer (Sec. 19.1) with the addition of the following:

(*a*) An entrance slit, which usually acts as the field stop of the collecting optics.

(*b*) A collimator, which may be a lens or a mirror, with the entrance slit at its focus.

(*c*) A dispersing element, either a prism or grating.

(*d*) A focusing element, which produces an image of the entrance slit from the parallel beam at each wavelength so that these images are dispersed linearly to form the familiar spectrum.

(*e*) One or more exit slits to select the radiation in any desired region of the spectrum and allow it to pass on to the detector.

A spectroradiometer can be said to be a radiometer with a monochromator located between the collecting optics and the detector.

19.3. Monochromators [4]

In self-collimating spectrometers and monochromators, a single optical element serves both for collimating the beam from the entrance slit and for refocusing the dispersed beam onto the plane of the exit slit. The most frequent configurations of self-collimating instruments is the Littrow type shown in Fig. 19-5. In this instrument, a plane mirror is positioned so as to return the dispersed beam back to the prism or

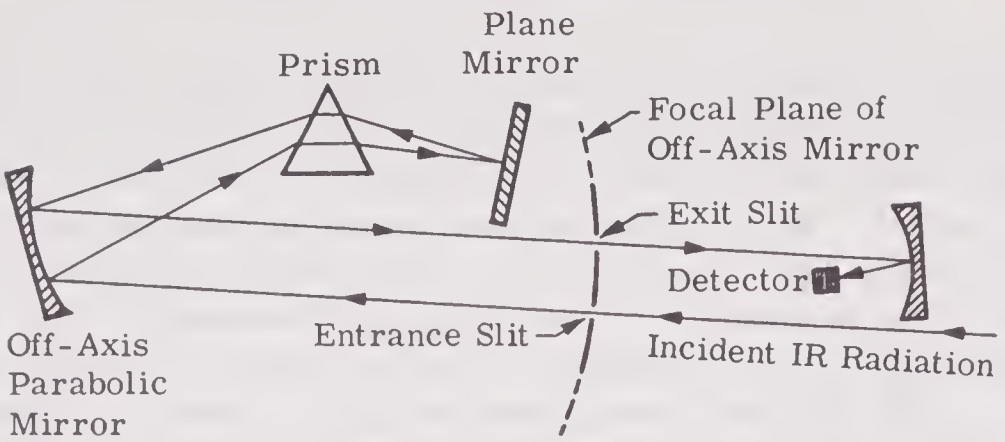


FIG. 19-5. Single-pass monochromator.

grating for a second dispersion before it returns to the self-collimating mirror, or lens, and the exit slit.

This type of monochromator has a limited resolution, determined by the finite size and quantity of the prism employed. An impure spectrum is also produced, particularly at the wavelengths greater than 8μ, because of the high-intensity short-wavelength light from the source being scattered into the exit slit. Various devices can be employed to improve purity of spectrum, such as the use of selective choppers or shutters which are transparent to the short-wavelength radiation, and thus only chop the desired longer wavelengths. The best solution can be achieved by using a double monochromator. If it is a Littrow type, such as the Perkin-Elmer Model 99 monochromator shown in Fig. 19-6, the beam is acted upon four times by the prism or grating. Not only is dispersion increased by these repeated dispersions but spectral resolution and purity are also improved. In a double monochromator, two monochromators are placed in series, with the exit slit of one forming the entrance slit of the second. In this configuration, the repeated dispersion also improves the spectral resolution and purity. The emerging beam from the exit slit of a well-designed double monochromator contains a minimum of scattered radiation of wavelengths outside the desired passband.

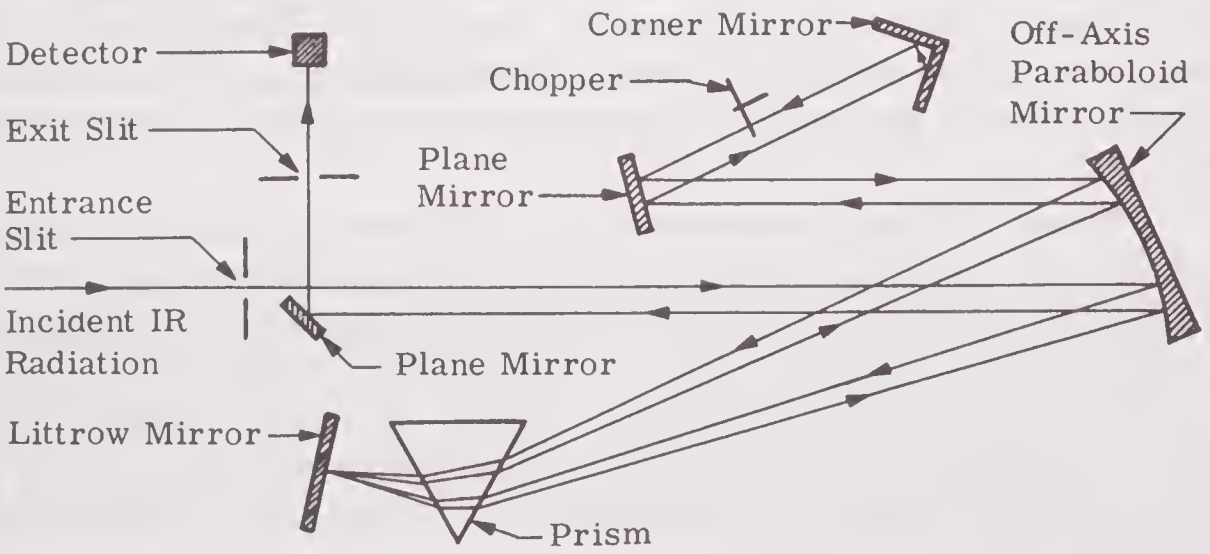


FIG. 19-6. Perkin-Elmer Model 99 double-pass monochromator.

In double-pass instruments, the second pass is made over portions of the same path as the first. In one type, the second-pass radiation is distinguished from the first, also emerging from the exit slit, only by the insertion of a chopper in a portion of the second-pass beam that does not overlap the first. In this case, although the modulated

beam from the second pass may have high spectral purity, a fairly high level of unchopped first-pass radiation of unwanted wavelengths can also be incident on the detector. The effect of this first-pass radiation on the response of the detector element to the chopped second-pass radiation of the desired wavelengths should be carefully tested in such cases. This is particularly important for measurements in wavelength regions in which the spectral responsivity of the detector element can be very low as compared to its responsivity to the unchopped wavelengths from the first pass. Spectral filters inserted at the entrance slit can be used to produce substantial changes in the level of this unwanted, unmodulated radiation in order to observe the effect, if any, on the output. Ideally, there should be no effect, because the output should be a measure only of the chopped radiation in the desired wavelength band. However, this needs to be verified for each detector element used over the complete range of wavelengths for which it will be employed.

The wavelength band in the output of a prism monochromator is a direct function of the relative positions of the slits, prism, and any optical elements (such as a Littrow mirror) used to shift the dispersed spectrum across the exit slit. A grating monochromator, however, can have overlapping orders. A fairly low-dispersion foreprism monochromator, or suitable filters, can be used in front of the entrance slit to remove wavelengths of the undesired orders and to eliminate the ambiguities.

19.4. Spectrometers and Spectrophotometers

An infrared spectrometer permits selection and isolation of a desired wavelength (or band of wavelengths) in the infrared spectrum for study of the physical properties of a material. An infrared spectrophotometer permits selection and isolation of a desired wavelength (or band of wavelengths) in the infrared spectrum, for simultaneous comparative examination of the physical properties of a sample material and a reference material at a selected point in the spectrum.

In spectrometers, single-beam photometry is used for radiation measurement. The radiation spectrum from the source alone is first measured; then a sample is introduced in the sample area. The source radiation, is modified by the sample, and is then measured. A comparison of these two measurements, as a function of radiation frequency, is then calculated and replotted. In spectrophotometers, double-beam optical systems are used in which radiant intensity through a sample cell and a reference cell is compared and automatically recorded. In spectrometers and spectrophotometers, the spectrum is usually dispersed by a monochromator employing a refraction prism or a diffraction grating.

19.4.1. Infrared Grating Spectrometry [6]. Infrared grating spectrometers are used for studies of molecular structure, where rapid automatic recording and extremely high resolution are required. When used with multiple reflection cells, long path-length studies in gases and liquids can be made. Fast recording is achieved by using highly sensitive lead sulfide, lead telluride, lead selenide, or indium antimonide detector cells with rapid response times. Improved gratings and recent improvements in detectors and in amplifier design have made possible the construction of instruments with resolving powers exceeding 150,000.

Optical systems are generally of the off-axis parabolic mirror type, or the on-axis Pfund type. Mirrors of high quality are used to obtain good images and high resolving power. The Pfund type employs on-axis paraboloid mirrors and plane mirrors with an aperture in the center to produce superior images and higher resolving power.

19.4.1.1. *Off-Axis Parabolic Grating Spectrometer* [7]. An off-axis type, double-pass instrument is shown in Fig. 19-7. It has a focal length of 10 m, achieved by mounting the grating in the fashion of the Littrow instrument to ensure double passage

of the radiation beam through the optical system. Using a grating of great perfection, a resolving power of 120,000 to 140,000 lines per inch in the 1.3- to 1.7-μ region is obtained, with an increase in resolving power of about one order of magnitude at longer wavelengths.

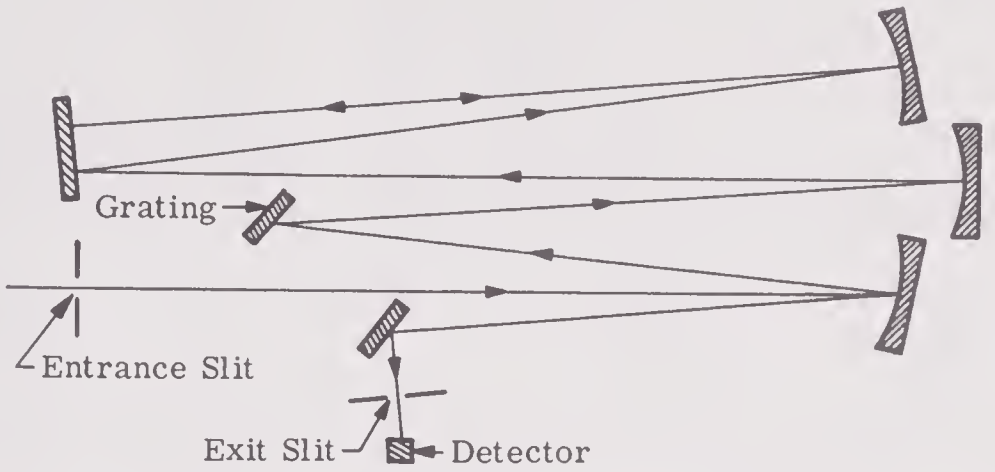


FIG. 19-7. Off-axis, double-pass grating spectrograph.

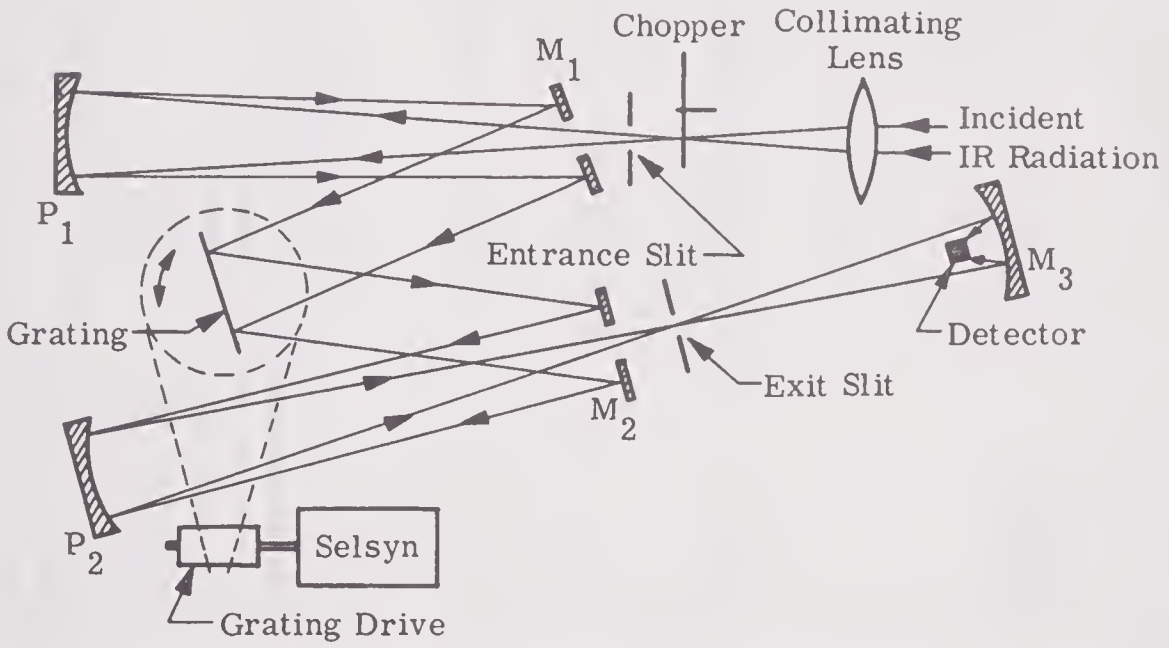


FIG. 19-8. On-axis Pfund grating spectrograph.

19.4.1.2. *On-Axis Pfund Grating Spectrometer* [6]. A typical on-axis Pfund-type direct-recording spectrometer is shown in Fig. 19-8. Incident infrared radiation focused by a collimating lens on the entrance slit and modulated by a chopper passes through the central aperture of plane mirror M_1. Reflected by the paraboloidal mirror, P_1, it emerges as a parallel beam of radiation, which is reflected by mirror M_1 to the grating. The grating is accurately located on a turntable, which may be rotated in order to scan the spectrum. From the grating, the diffracted beam, reflected by mirror M_2, is focused by a second paraboloid, P_2, through the central aperture of mirror M_2 to the exit slit. The emerging beam is then focused by the ellipsoidal mirror, M_3, on the detector. This type of on-axis system produces a better spectral image and superior resolution.

19.4.2. Commercially Available Spectrometers and Spectrophotometers. Table 19-3 lists some of the characteristics of some commercially available spectrometers and spectrophotometers. Representative instrument types are described in the following paragraphs.

TABLE 19-3. CHARACTERISTICS OF COMMERCIALLY AVAILABLE SPECTROMETERS AND SPECTROPHOTOMETERS

SPECTROMETERS	*Monochromator*		*Operation*	*Resolution*	*Spectral Range* (μ)	*Scan Speed (total scan)*	*Accuracy*	*Wavelength Reproducibility*
	Type	(f/no)						
Perkin-Elmer Corp. 112	Double pass Littrow, NaCl prism	4.5	Single beam	$<0.01\ \mu$ at $12\ \mu$	2-15	22 min	–	$0.01\ \mu$ at $2\ \mu$ $0.003\ \mu$ at $15\ \mu$
Rapid-Scan	Rapid-scan, double pass nutating mirror functioning as Littrow mirror; fused silica or NaCl prism	5	–	–	0.20-13.5	–	–	–
Block Engineering, Inc. I-4 s	–	–	–	40 cm^{-1} max.	0.35-3	–	–	–
I4T	–	–	–	40 cm^{-1} max.	1-16	–	–	–
Barnes Eng. Co. Multichannel 15-140	–	–	–	$0.1\ \mu$	1.3-2.4	–	–	–

SPECTROPHOTOMETERS

Perkin-Elmer Corp. 13-U	Single-pass Littrow mounting, NaCl prism	4.5	Split double-beam direct-ratio phase discrimination; full aperture single beam, direct intensity	0.2 mμ at 250 mμ 1.5 mμ at 580 mμ 5.0 mμ at 2.0 μ 0.02 μ at 10 μ	190-50	–	–	–
KBr 137	70° prism, KBr	5	Double beam, optical null	–	12.5-25	–	$\pm0.03\ \mu$ $\pm1\%$	$\pm0.01\ \mu$ $\pm1\%$
137-B	70° prism, NaCl	5	Double beam, optical null	–	2.5-15	–	$\pm0.03\ \mu$ $\pm0.1\%$	$\pm0.01\ \mu$ $\pm0.5\%$

TABLE 19-3. CHARACTERISTICS OF COMMERCIALLY AVAILABLE SPECTROMETERS AND SPECTROPHOTOMETERS (*Continued*)

SPECTROPHOTOMETERS	*Monochromator*		*Operation*	*Resolution*	*Spectral Range* (μ)	*Scan Speed* (*total scan*)	*Accuracy*	*Wavelength Reproducibility*
	Type	(f/no)						
137-G	Double grating	5	Double beam, optical null	4 cm^{-1} beyond 4 μ 10 cm^{-1} 1.6 μ 10 cm	0.83-7.65	8 or 24 min	$\pm$0.004 μ $\pm$1% $\pm$0.012 μ $\pm$1%	$\pm$0.0015 μ $\pm$0.5% $\pm$0.004 μ $\pm$0.5%
202	Littrow mirror, 60° fused silica prism	5	Double beam, optical null	0.2 mμ at 210 mμ 0.4 mμ at 390 mμ 1.5 mμ at 600 mμ	0.190-0.390 0.350-0.750	2 or 8 min	$\pm$0.5 μ in UV $\pm$1 μ in visible	0.3 μ in UV 0.5 μ in visible
237	Double grating	5	Double beam, optical null	–	2.5-7.7 5-16	8 and 24 min	–	–
Far Infrared 301	Model 210	3.8	Double beam	0.7 cm^{-1}	660-60 cm^{-1}	–	–	–
350	Double Littrow monochromator, 30° fused silica prisms	9	Double beam	0.03 mμ at 185 mμ 1.5 mμ at 2.6 μ	0.175-2.7	15 sec to 10 min	0.03 μ at 175 μ 2.0 μ at 700 μ 2.2 μ at 2.7 μ	0.04 μ at 200 μ 1.5 μ at 700 μ 3.0 μ at 2 μ
421	Single-pass Littrow mount with 2 gratings	4.5	Double beam, optical null	Better than 0.3 cm^{-1} at 2200 cm^{-1}	2.5-18.2	77 cm^{-1} sec^{-1} to 0.22 cm^{-1} min^{-1}	1 cm^{-1}	0.5 cm^{-1}
Portable	–	2.7	–	Less than 330 A (with 1-mμ slit)	2.2-μ band	–	–	–
Beckman Instruments, Inc.,								
IR-4	Double monochromator, 60 × 75 mm NaCl prism	–	Double beam, optical null; Single beam, linear amplifier	0.01 μ at 10 μ	1-15	Variable	0.015 μ	0.005 μ
IR-5A (NaCl)	Single monochromator, 66° prism, NaCl	–	Double beam, optical null; Single beam, electrical null	2.1 cm^{-1} at 911 cm^{-1}	2-16	3 min 15 min	0.025 μ	0.01 μ

TABLE 19-3. CHARACTERISTICS OF COMMERCIALLY AVAILABLE SPECTROMETERS AND SPECTROPHOTOMETERS (*Continued*)

SPECTROPHOTOMETERS	*Monochromator*		*Operation*	*Resolution*	*Spectral Range* (μ)	*Scan Speed* (*total scan*)	*Accuracy*	*Wavelength Reproducibility*
	Type	(f/no)						
IR-5A (CsBr)	Single monochromator, 66° prism, NaCl	–	Double beam, optical null; Single beam, electrical null	Better than 4 cm^{-1} at 370 cm^{-1}	11-35	3 min 15 min	0.04 μ	0.01 μ
IR-7	Prism-grating double monochromator, 66° prism, NaCl	10	Double beam, optical null; Single beam, linear amplifier	0.30 cm^{-1} at 1000 cm^{-1}	2.5-15.4	–	0.5 cm^{-1}, 1st order 2.0 cm^{-1}, 2nd order 2.5 cm^{-1}, 3rd order 5 cm^{-1}, 4th order	0.25 cm^{-1} in 1st and 2nd orders 1.0 cm^{-1} in 3rd and 4th orders
Carry-White Model 90	Prism grating double monochromator, 66° KBr prism, Littrow mirror	–	Double beam	0.25 cm^{-1} or better	4000-450 cm^{-1}	2 min to 1000 hr	0.5 cm^{-1}	0.1 cm^{-1}
Baird Atomic, Inc. NK-1	100 × 80 mm, 60° prism	7.5	Double beam, optical null; Single beam optional	0.015 μ from 2 to 16 μ	0.25-38	0.1 min μ^{-1} to 1 hr μ^{-1}	±0.015 μ ±0.5%	±0.005 μ ±0.1%
NK-3	Modified Czerny-Turner prism grating	–	Double beam, optical null; Single beam optional	0.25 cm^{-1} at 950 cm^{-1}	2-22	20 min	±4 cm^{-1} at 5000 cm^{-1} ±0.5%	±0.5 cm^{-1} at 950 cm^{-1} ±5% or better

19.4.2.1. *Single-Beam Double-Pass Spectrometer.* An example of a single-beam double-pass infrared spectrometer is the Perkin-Elmer Model 112 instrument shown in Fig. 19-9. Infrared radiation from a source is focused by mirrors M_1 and M_2 on the entrance slit, S_1, of the monochromator. The radiation beam from S_1, path 1, is collimated by the off-axis paraboloid, M_3, and a parallel beam traverses the prism for a first refraction. The beam is reflected by the Littrow mirror, M_4, through the prism for a second refraction, and focused by the paraboloid, path 2, at the corner mirror, M_6. The radiation returns along path 3, traverses the prism again, and is returned back along path 4 for reflection by mirror M_7 to the exit slit, S_2. By this double dispersion the radiation is spread out along the plane of S_2. The frequency interval which passes through S_2 is focused by mirrors M_8 and M_9 on the thermocouple, TC. The beam is chopped by CH, near M_6, to produce an ac voltage (at the thermocouple) which is proportional to the radiant power or intensity of the beam. This voltage is amplified and recorded by an electronic potentiometer. Motor-driven rotation of Littrow mirror M_4 causes the infrared spectrum to pass across exit slit S_2, permitting measurement of the radiant intensity of successive frequencies.

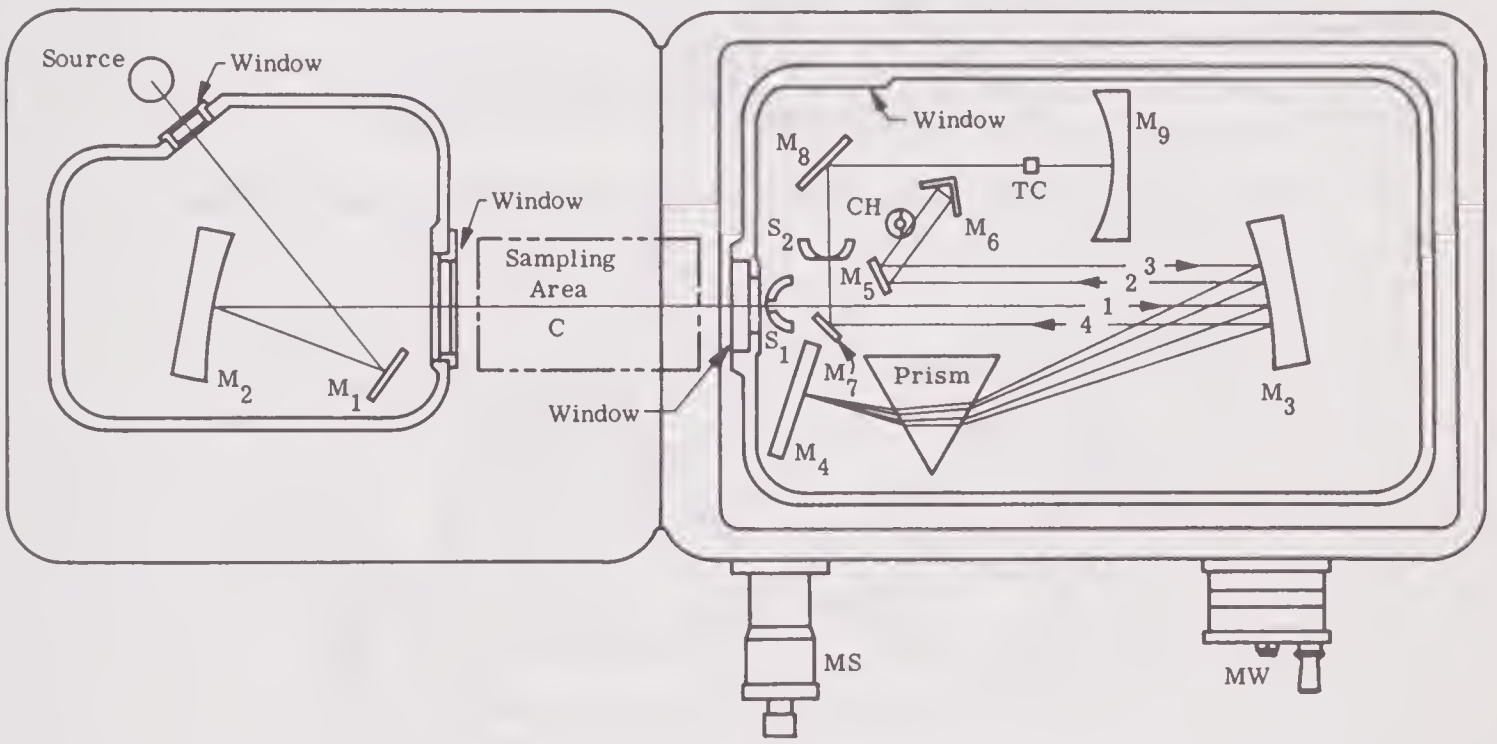


FIG. 19-9. Perkin-Elmer Model 112 single-beam double-pass infrared spectrometer.

19.4.2.2. *Prism-Grating Double Monochromator Spectrometer* [6]. This instrument, manufactured by Unicam Instruments, Ltd., uses a prism-grating double monochromator covering the widest possible spectral range without breaks for optical changes or adjustments. The instrument is shown in Fig. 19-10.

The prism monochromator has four interchangeable prisms, and the grating monochromator has two interchangeable gratings. The two monochromators, gaged by cams linear in wave number, are driven by a common shaft. The instrument may be used either as a prism-grating double monochromator, or as a prism spectrometer by blanking the grating monochromator. Gratings, prisms, and cams may be automatically interchanged by means of pushbuttons. Magnetically operated slits, programmed by a tapped potentiometer, provide a constant energy background. A star wheel time-sharing beam attenuator is used in the double-beam photometer.

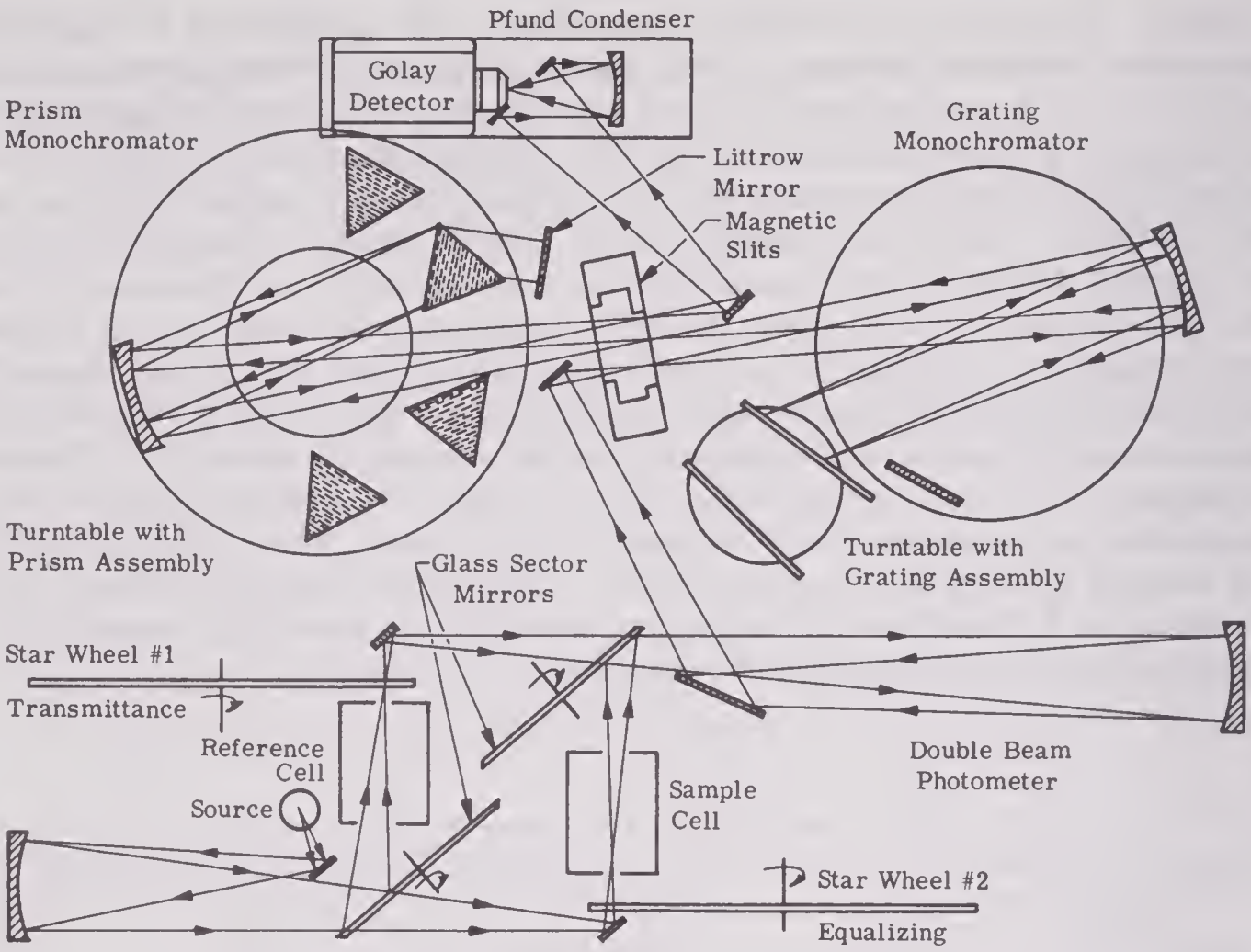


FIG. 19-10. Unicam prism-grating double monochromator spectrometer.

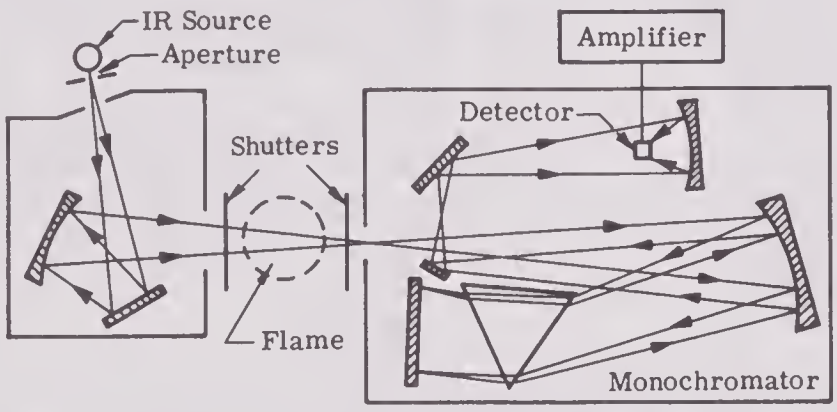


FIG. 19-11. Infrared flame-temperature spectrometer.

19.4.2.3. *Flame-Temperature Spectrometer* [6]. This spectrometer provides a method to measure the temperatures of missile flames up to 5 ft in diameter up to several thousand degrees centigrade. The instrument, shown in Fig. 19-11, is rapid and accurate, and requires no calibration or attenuation of the gas stream.

Infrared radiant energy from a source of known emission passes through the hot gas stream with absorption bands due to the water vapor and carbon dioxide present in the gas stream as products of combustion. The infrared beam is focused on the entrance slit of a Perkin-Elmer Model 98 monochromator. At a given wavelength, with the shutter out of the beam, the radiant energy E_1 is measured:

$$E_1 = E_s\tau + E_f(1 - \tau) \tag{19-13}$$

where E_s = known energy from the source

E_t = energy from blackbody source at gas temperature T

τ = unknown percent transmission of carbon dioxide and water vapor in gas stream

With the shutter in the beam, the energy E_2 is measured:

$$E_2 = E_t\,(1-\tau) \tag{19-14}$$

The two equations are then solved for τ and E, and the apparent temperature of the gas stream is obtained.

19.4.2.4. *Rapid-Scan Spectrometer.* The rapid-scan spectrometer manufactured by Perkin-Elmer Corp. (Fig. 19-12) records the distribution and time variation of the spectral wavelengths of the energy radiated during the powered flight portions of missile firings. The instrument consists of a rapid-scan monochromator, a radiation detection system, and appropriate readout and recording equipment.

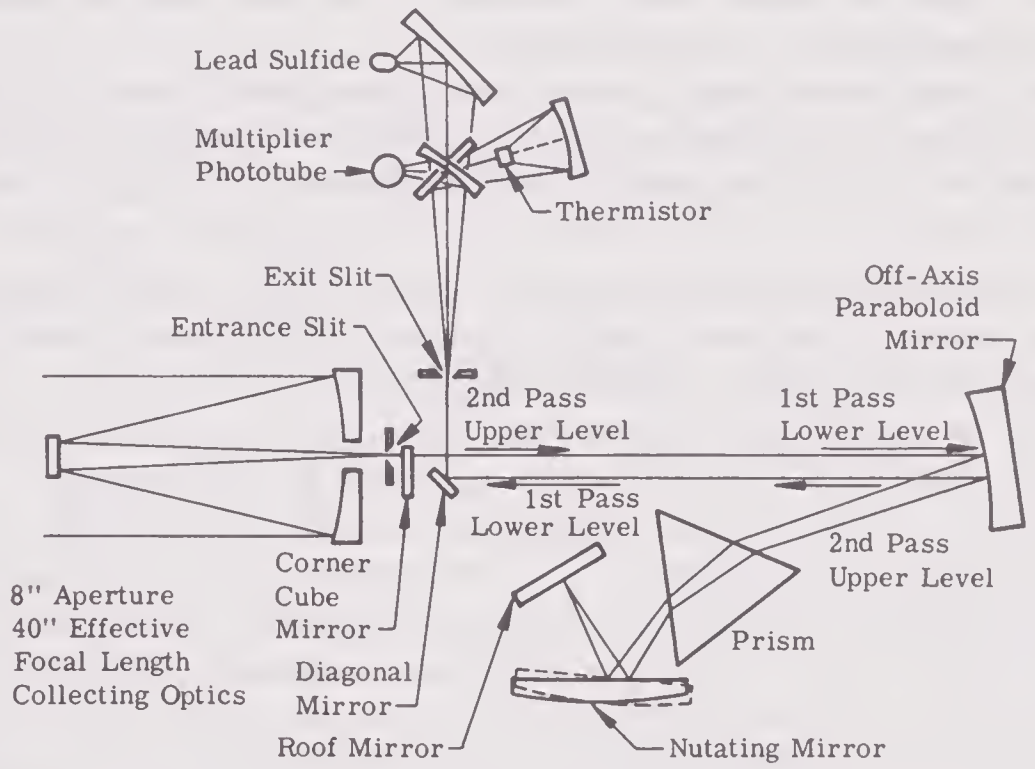


FIG. 19-12. Perkin-Elmer rapid-scan spectrometer.

The rapid-scan monochromator utilizes a double-pass system where the first and second passes are physically separated. As shown in Fig. 19-12, radiant energy from the source is imaged by the collecting optics, composed of a modified double-pass Littrow system, on the entrance slit of the monochromator. The beam is collimated by an off-axis paraboloid mirror onto the prism. The energy beam is then refracted to the nutating mirror, reflected to the roof mirror (where the vertical light motion component is eliminated), and then back through the prism system. The returned beam is brought to a focus by the paraboloid mirror at the corner cube mirror, where it is displaced up and across, back to the paraboloid mirror. It then passes through the monochromator for the second time. The second pass is 1 in. higher than the first pass, and the second pass only is intercepted by the diagonal mirror and focused on the exit slit.

The detector sees an alternating signal whose frequency is the chopped frequency; within each half cycle the signal amplitude varies in accordance with the radiant spectral energy. The result is a plot of the spectrum of the source repeated at the chopping frequency.

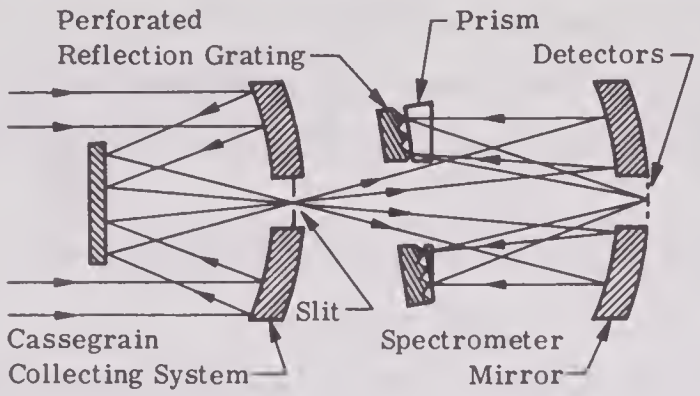


FIG. 19-13. Barnes in-line spectrometer.

19.4.2.5. *In-Line Spectrometer.* The lens configuration in the Barnes instrument (Fig. 19-13) permits effective use of noncollimated light with reflection gratings. The major advantage of using converging light in spectrometers is that it permits construction of an in-line instrument. In this technique, light rays received at one end travel straight through the spectrometer. The rays are then filtered into narrow wave bands and are dispersed into a useful spectrum. They may then be photographically or electronically recorded at the opposite end.

The in-line spectrometer uses a correcting prism placed ahead of the refraction grating. The prism is designed so that its differential refracting effect on rays of various inclination exactly compensates for the difference in light-path length. Thus, all rays contributing to the resulting spectrum line will focus in one plane.

19.4.2.6. *Interferometer Spectrometer.* The Block Engineering interferometer-spectrometer has greater throughput than conventional spectrometer, because the interferometer has a large entrance aperture determined by the mirror size. This enables the instrument to accept more radiant energy from the source than prism or grating instruments in which the entrance aperture is limited by narrow slits.

High sensitivity gain is due to the instrument's examining each wavelength throughout the entire time period of each scan. In a conventional instrument, each wavelength is examined for only a very short part of the scan time ($1/n$th the scan time if n is the number of resolution elements); the interferometer achieves a gain which is $\sqrt{n}$ for the same scan time. For typical instruments this can be a factor of 50. Furthermore, this gain is realized even when one examines point sources where the throughput gain is not large.

A block diagram of an interferometer spectrometer is shown in Fig. 19-14. Incoming infrared radiation is received by the interferometer, and a fringe pattern is produced. When one of the mirrors in the interferometer is moved back and forth at a slow, constant velocity, the motion is manifested as an alternate brightening and darkening of the central fringe.

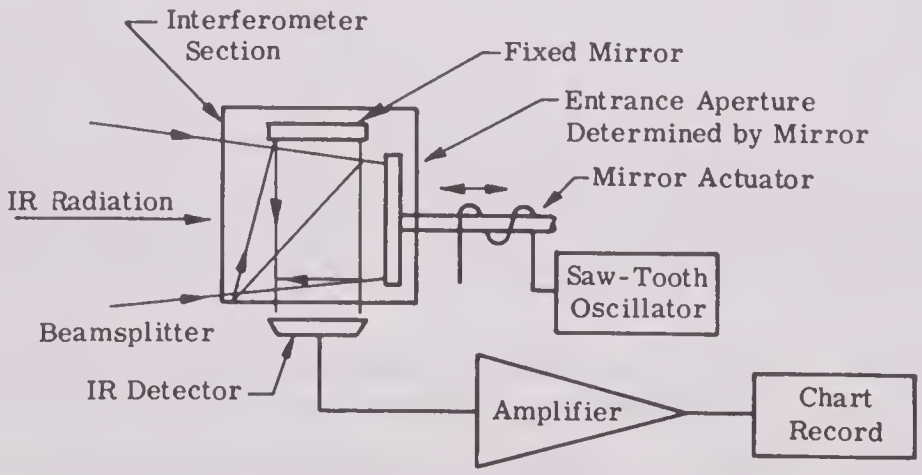


FIG. 19-14. Block Associates interferometer spectrometer.

An infrared detector placed at the central fringe converts these cyclic changes into an ac signal. If the mirror velocity is kept constant at a predetermined value, the frequency of the ac signal from the detector is directly related to the wavelength of incident radiation, assuming that the incident radiation is at one given wavelength (monochromatic).

If another wavelength twice as long as the first (half the frequency) should be substituted as the incident radiation source, the ac output signal from the detector would be at one-half the frequency of the first. The amplitudes of the two signals would remain the same if the maximum brightness of the two sources were the same.

If incident radiation containing many wavelengths were introduced into the system, the output of the detector would consist of a superposition of all the ac signals which correspond to all the wavelengths in the source.

The output of the interferometer system is tape recorded and played back through an audio wave analyzer to recover the infrared spectrum.

The scan drive must be linear and constant and the effects of source-intensity variations must be negligible. Either of these can cause apparent spectral peaks in the output.

19.4.2.7. *Double-Beam Optical Null Spectrophotometer.* An example of a double-beam optical null system is the Beckman spectrophotometer shown in Fig. 19-15. The instrument utilizes a double monochromator which is convertible to single-beam operation.

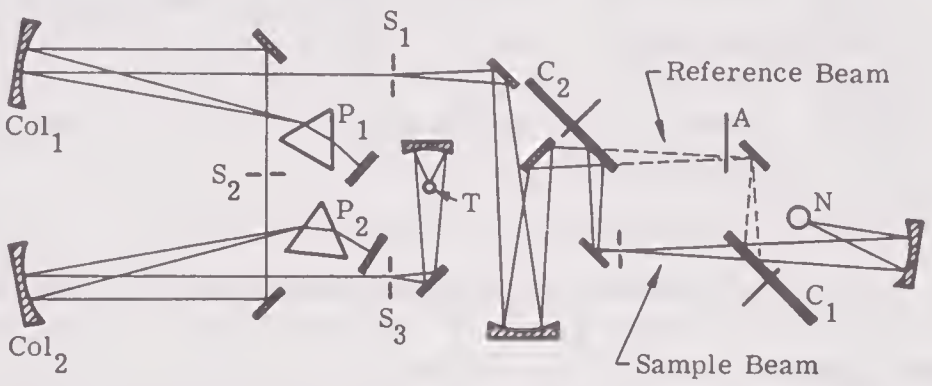


FIG. 19-15. Beckman IR-4 automatic recording infrared spectrophotometer.

For double-beam operation, radiation from a Nernst glower, N, passes through the half mirror, C_1, rotating at 11 cps. The radiation then passes through the sample during one half revolution of the mirror, where it is recombined by the symchronously rotating half mirror, C_2. During the other half revolution of the mirror, the beam is deflected through the reference. Mirror C_1 chops the radiation beam at 11 cps and directs it alternately through the sample and through the reference.

The recombined beam is then directed through the controllable entrance slit, S_1, into the first of two monochromators. In the first monochromator, the beam is transmitted by a collimator, Col_1, through a prism, P_1, and associated Littrow mirrors for initial dispersion. This dispersed beam then passes through fixed-width entrance slit S_2, into the second monochromator. Here collimator Col_2 transmits it through prism P_2 and Littrow mirrors for further dispersion.

The doubly dispersed monochromatic beam now exits through the controllable exit slit, S_3, and is focused on a thermocouple, T. The signal from the thermocouple is amplified and used to position an optical attenuator, A, in the beam path, so that the radiation transmitted by sample and reference beams are equal in intensity. The position of the optical attenuator determines the position of the recorder pen, producing linear wavelength records.

For single-beam operation, the reference beam is blocked. The thermocouple receives only energy from the sample beam, chopped at 11 cps. The amplified thermocouple output is recorded directly by the pen.

19.4.2.8. *Direct-Ratio Spectrophotometer.* A direct-ratio system is used in the Perkin-Elmer instrument (Fig. 19-16). This is the same as the optical null system in that the two radiation beams are simultaneously compared. In this case, however, a separate electrical signal proportional to each beam is separated by a phase-sensitive detection signal. The reference beam is placed across the slide wire of a recording potentiometer and the sample signal fed to a standard strip chart recorder so that the ratio between reference and sample beams is automatically recorded. The optical system of the instrument is shown in the figure.

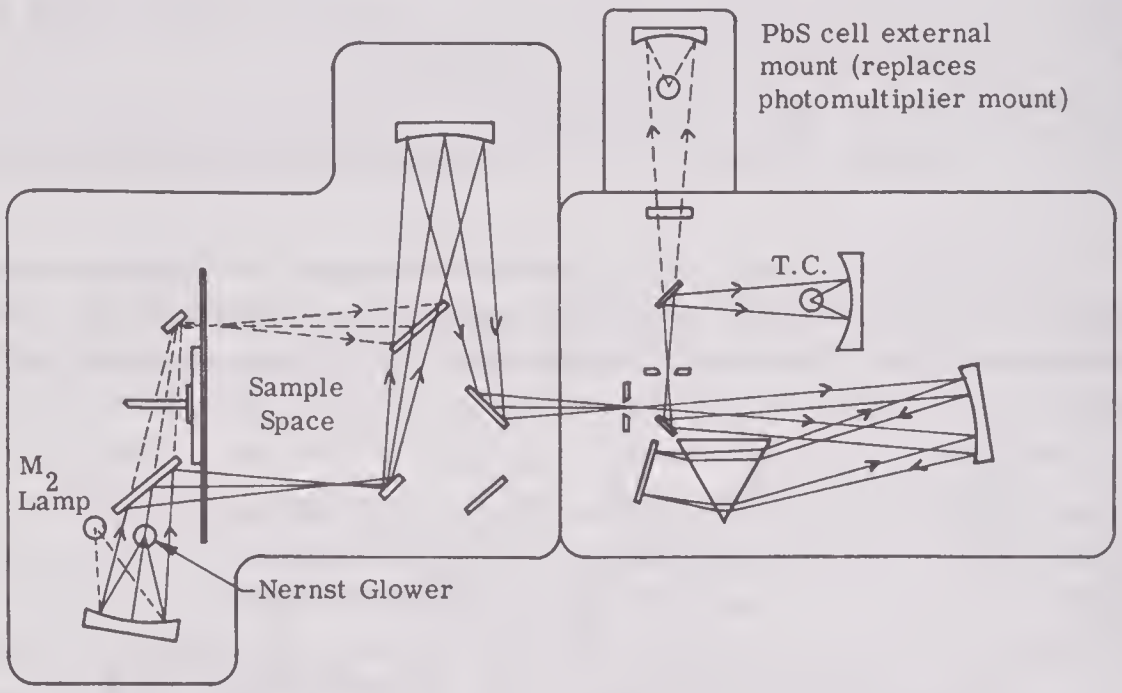


FIG. 19-16. Perkin-Elmer universal spectrophotometer Model 13-U.

19.5. Interferometers

Interferometers divide a beam of light into two or more parts which travel different paths and recombine to form interference rings. The form of these rings is determined by the difference between the optical paths of the successive beams. Interferometers measure the difference in optical path length and refractive index.

19.5.1. Rayleigh Interferometer [7]. In the Rayleigh interferometer shown in Fig. 19-17, monochromatic light from linear source S_1 falls on a screen and is split into two beams by well-separated slits, S_2 and S_3. The light is rendered parallel by lens L_1 and passes through two exactly similar tubes, closed at each end by transparent windows. The contents of the tubes (liquid or gaseous) determine the positions of the fringes by the path difference introduced between the two beams. The beams leaving the tubes are focused by lens L_2, producing an image whose bandwidth is determined by the width of S_2 and S_3 and by the magnification of the system.

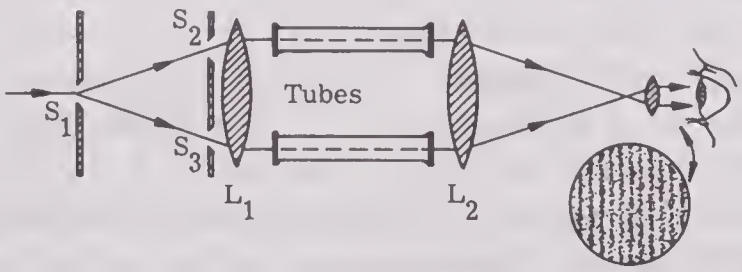


FIG. 19-17. Rayleigh interferometer.

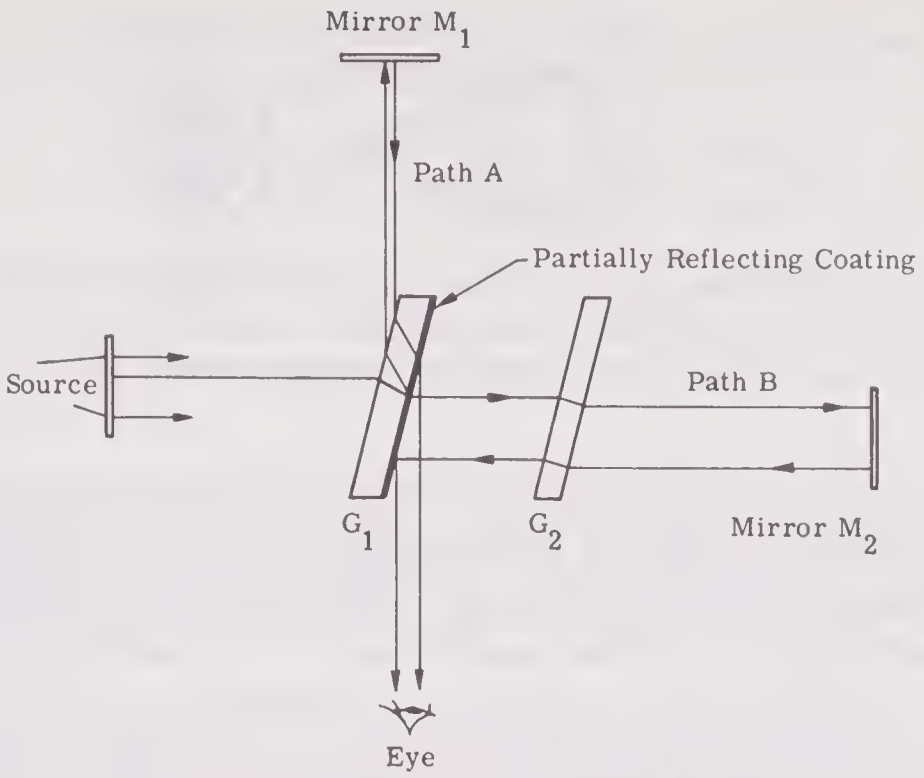


FIG. 19-18. Michelson interferometer.

19.5.2. Michelson Interferometer. The Michelson interferometer, shown in Fig. 19-18, consists of two plane mirrors, M_1 and M_2, one of which is adjustable, and two plane-parallel plates, G_1 and G_2. Light from an extended source is incident at 45° on plate G_1, (partially silvered on the rear surface) and is divided into reflected (path A) and transmitted beams (path B) of equal intensity. The light reflected from mirror M_1 passes through plate G_1 a third time before it reaches the eye. The light reflected from mirror M_2 passes back through G_2 a second time, is reflected from the surface of plate G_1, and into the eye. The two beams have a phase difference governed by the difference in the two paths.

Compensating plate G_2 compensates for the passage of light through the plate in path A. Its use is not essential for producing fringes in monochromatic light but is indispensable when white light is used. The light from every point on the extended source interferes with itself according to the distance between mirrors, or according to the different length of arms A and B. Constructive interference will occur when

$$2d \cos \theta = m\lambda \tag{19-15}$$

where d = path difference
θ = angle to a source element imaged by M_1 or M_2 as seen by the eye

19.5.3. Twyman-Green Interferometer [7]. In the Twyman-Green interferometer, a Michelson interferometer is illuminated with strictly parallel monochromatic light, produced by a point source at the principal focus of a well-corrected lens. As shown in Fig. 19-19, light comes from a pinhole, P, at the principal focus of a lens, L_1. A second lens, L_2, focuses the emerging light into the eye. By the use of collimated light, the fringes at infinity can be seen at a finite distance (such as on one of the surfaces of a prism) because of the greatly increased depth of focus. In addition, the light is made to traverse the optical part under test (a prism in Fig. 19-19), making the result of the test explicit.

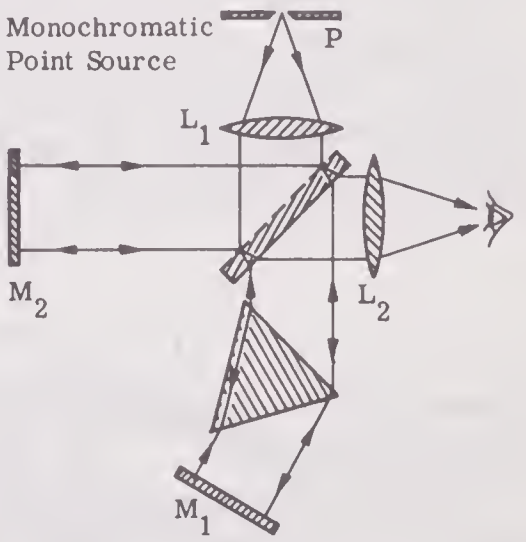


FIG. 19-19. Twyman-Green interferometer.

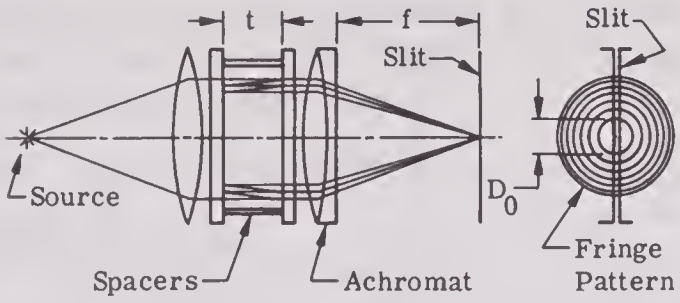


FIG. 19-20. Fabry-Perot interferometer.

19.5.4. Fabry-Perot Interferometer [8]. The Fabry-Perot interferometer, shown in Fig. 19-20, consists of two quartz or glass plates, one face of each worked very flat (1/20*th*-1/100*th* fringe), partially aluminized, and separated by a spacer which renders the two faces exactly parallel. The outer faces form a small wedge angle with the working faces to prevent spurious fringes. Any entering ray will be reflected back and forth between the faces, and at each reflection some radiation will be transmitted. Some five to thirty successive rays which result from a single incident ray will be transmitted. These reinforce each other if the path distances between the two internal faces are integral multiples of the wavelength. The condition for obtaining a bright fringe is given by $n\lambda = 2t \cos\theta = 2t$, approximately, since θ is small. The pattern projected on the slit of the spectrograph is a series of concentric, circular interference fringes governed by the relation

$$n = \frac{2t}{\lambda} + \frac{2t}{\lambda}\left(\frac{D^2}{8f^2}\right) \tag{19-16}$$

where n = order of interference at the center of the ring system
t = separation of the plates
D = diameter of a particular ring of wavelength
f = focal length of the lens which projects the fringes on the slit

The term $2t/\lambda$ represents the integral order of interference of the innermost ring, and the last term of Eq. (19-16) represents the fractional order between a ring of diameter D and the center.

The quantity t must first be determined by measurements on a known wavelength standard. Thereafter the wavelengths of unknown lines can be measured to thousandths of an angstrom, if the lines themselves are sufficiently sharp.

As the thickness of the aluminized or multiple-layer coating increases, the number of the reflections, and hence of the interfering rays, increases, so that the fringes become sharper. Absorption also becomes greater, causing the intensity of the pattern to decrease. Reflectivity is generally considered the best compromise between high resolution and low intensity. Quantitatively, the resolving power is given by

$$\frac{\lambda}{d\lambda} = \frac{2\pi n}{\sqrt{4.45}} \frac{\sqrt{r}}{(1-r)} \sim n \cdot \frac{3\sqrt{r}}{1-r} \tag{19-17}$$

where $3\sqrt{r}/(1-r)$ can be considered physically as a measure of the effective number of interfering beams.

The Fabry-Perot interferometer achieves the highest resolution of any known optical element. Wavelength measurements can be made to thousandths, or even ten thousandths, of an angstrom, and the instrument is used almost exclusively for ultrahigh resolution.

For a Fabry-Perot interferometer, the free spectral range, expressed in angstroms, is given by $F_\lambda = \lambda^2/2t$, so that for Hg (= 5461 A),

$$F_\lambda = \frac{(5461)^2}{2 \times 10^8} = 0.15 \text{ A} \tag{19-18}$$

Thus, a high-dispersion high-resolution spectrograph is required to avoid overlapping orders in the Fabry-Perot pattern. In practice, a diameter of the ring pattern of the etalon is focused on the slit of a large spectrograph and some overlapping is tolerated for lines closer than 1 to 5 A.

19.5.5. Spherical Fabry-Perot Interferometer [9]. At resolutions of several million, the luminosity of the Fabry-Perot interferometer decreases and sources at resolutions in this order are inherently weak. The spherical Fabry-Perot interferometer permits considerable progress in the study of phenomena in the area of high resolution.

The instrument, shown in Fig. 19-21, consists of two spherical mirrors whose separation is equal to the radius of curvature r, so that the paraxial foci coincide and the instrument is an afocal system. One half of the surface of the mirrors is semi-reflecting and the other half is fully reflecting. Any incident ray gives rise to an infinite number of outgoing rays which are coincident, and not only parallel, as with the plane Fabry-Perot instrument. Neglecting aberrations, their path difference is $\Delta = 4r$, a constant. This requires that both mirrors be stopped down with circular diaphagms of diameter d. It can be shown that d increases with r, and that the light-gathering power is proportional to r (and thus to the theoretical resolving power R_0 instead of being inversely proportional to R_0). For practical useful values of r (a few centimeters), d always remains quite small (a few millimeters), permitting relatively easy figuring of nearly perfect spherical plates of this size.

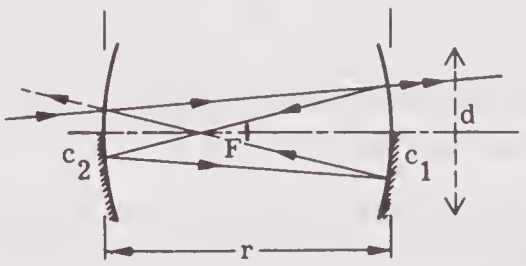


FIG. 19-21. Spherical Fabry-Perot interferometer.

19.5.6. Lummer-Gehrcke Plate. The Lummer-Gehrcke plate (Fig. 19-22) utilizes the interference between successive reflections in a thin quartz plate. Light is introduced to the plate by a total reflection prism, L_1. It then undergoes multiple internal reflections very near the critical angle of total reflection. The beams emerging at a grazing angle are brought to interference by a second lens, L_2. High reflectance and resolving power are thus obtained with unsilvered surfaces.

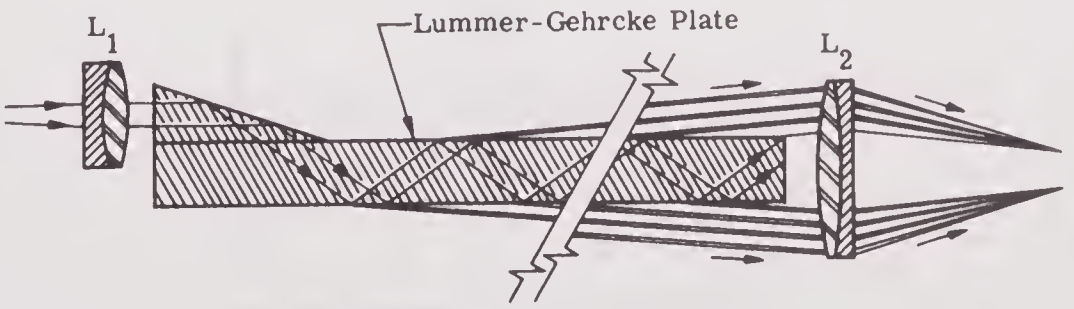


FIG. 19-22. Lummer-Gehrcke plate interferometer.

19.5.7. Spectral Transmittance of Interferometers [4]. The spectral transmittance functions, $\tau(\lambda)$ of the Fabry-Perot, Lummer-Gehrcke plate, and Michelson interferometers are described below.

Fabry-Perot interferometer:

$$\tau(\lambda) \sim \left(\frac{\eta^2/4\rho}{1 + \eta^2 \sin^2 \pi\rho}\right) \tag{19-19}$$

where $\eta^2 = \dfrac{4\rho}{(1-\rho)^2}$

ρ = the reflectance of the coatings

p = the order number

and

$$\Delta\lambda = \frac{\lambda^2}{2\pi t} \cdot \frac{1-\rho}{\rho} \tag{19-20}$$

where t = spacer thickness

Lummer-Gehrcke plate:

$$\tau(\lambda) \sim \frac{(1-\rho^N)^2 + 4\rho^N \sin^2(\pi\delta N/\lambda)}{(1-\rho)^2 + 4\rho \sin^2(\pi\delta/\lambda)} \tag{19-21}$$

where N = number of emergent beams obtained for a given plate length and angle of emergence

δ = path difference between two successive beams

Michelson interferometer:

$$\tau(\lambda) \sim F_1 \cdot F_2 \cdot F_3 \tag{19-22}$$

where F_1, F_2 are factors for the Fraunhofer patterns obtained with rectangular objects (they are of the $(\sin^2 \beta)/\beta^2$ form)

F_3 describes the interference pattern due to the N steps of the echelon

Figure 19-23 shows the spectral transmittance in a Michelson interferometer, where

a = amplitude of initial incident radiation at B

$a/2$ = amplitude of radiation in paths BM_1 and BM_2

a' = amplitude of exit radiation

M_1, M_2 = mirrors

B = beamsplitter

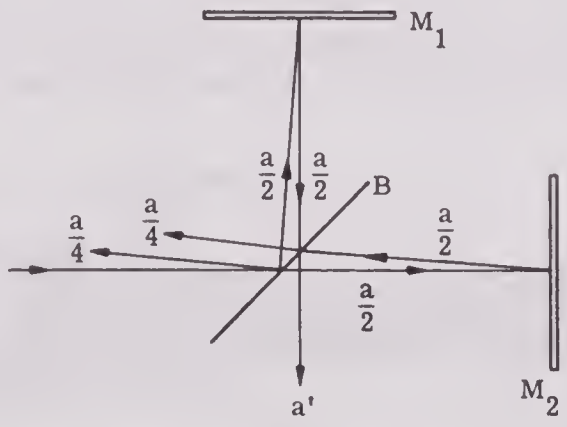


FIG. 19-23. Spectral transmittance in Michelson interferometer.

The amplitude, a, of an incident wave can be considered to be split equally into two parts by the beamsplitter, B, neglecting factors such as phase shift on reflection, the need for compensation plates, and other departures from the ideal case. Assuming M_1 and M_2 to be perpendicular, the divided amplitudes $a/2$ are associated with the two paths, BM_1 and BM_2; the amplitude of the emergent wave, a', is the result of the vector addition of one-half of each of these amplitudes. The emergent intensity distribution is obtained by $a'^* a'$, where a'^* is the complex conjugate of a'. Thus the transmittance for a Michelson interferometer for wave fronts which ara parallel to M_1 and M_2 is

$$\tau(\lambda) \sim a^2 \cos^2(\delta/2) \tag{19-23}$$

where $\delta = 4\pi t/\lambda$

t = relative displacement of M_1 from the zero phase position, *i.e.*, from $BM_1 = BM_2$

In the absence of multiple reflections, the interference fringes obtained are quite broad as compared to those from a Fabry-Perot interferometer. However, high resolution can still be obtained with long path differences if the problem of overlapping orders can be easily solved. If the movable mirror is oscillated in a sawtooth wave so that the mirror moves at a constant velocity for an appreciable portion of a wave cycle, then the transmission of the interferometer also varies for each wavelength. The output frequencies of such an interferometer-spectrometer are directly related to the wave numbers present in the incoming radiation. Wavelength identification is not difficult in this method of operation. On the other hand, Fabry-Perot interferometers used with large spacers require careful order sorting. Graphical means exist which can greatly simplify this problem.

References

1. M. R. Holter *et al.*, *Fundamentals of Infrared Technology*, Macmillan, New York (1962).
2. *Infrared Target and Background Radiometric Measurements*: *Concepts, Units, and Techniques*, Report of the Working Group on Infrared Backgrounds (WGIRB), Rept. No. 2389-64-T, Institute of Science and Technology, The University of Michigan, Ann Arbor, Mich. AD 289 810.
3. H. I. Sumnicht, "Infrared 'Effective' Radiation," *Proc. IRIS, IV*, 1, 52 (1959).
4. F. E. Nicodemus and G. J. Zissis, *Methods of Radiometric Calibration*, Rept. No. 4613-20-R, Institute of Science and Technology, The University of Michigan, Ann Arbor, Mich. (October 1962).
5. *Techniques*, Barnes Engineering Co., Stamford, Conn. (Spring 1956).
6. H. L. Hackforth, *Infrared Radiation*, McGraw-Hill, New York, (1960), Chapter 9.
7. J. Strong, *Concepts of Classical Optics*, Freeman, San Francisco (1958), Appendix B.
8. G. L. Clark, *The Encyclopedia of Spectroscopy*, Reinhold, New York (1960), pp. 258, 259.
9. P. Jacquinot, "New Developments in Interference Spectroscopy," *Repts. Progr. Phys.*, 23, 267-312 (1960).

Chapter 20
SPACECRAFT THERMAL DESIGN

L. H. Hemmerdinger and R. J. Hembach

Grumman Aircraft Engineering Corporation

CONTENTS

20.1. Definitions 784
20.1.1. Thermal Radiation Properties Terms 784
20.1.2. Space Technology Terms 785
20.2. Basic Heat-Transfer Theory 785
20.2.1. Conduction 785
20.2.2. Radiant Heat Transfer Between Two Surfaces 787
20.3. Thermal Coatings for Spacecraft 787
20.3.1. Optical Design 787
20.3.2. Environmental Consideration 792
20.4. Solar System 793
20.5. Angular Dependence of Reflectivity 793
20.6. Measurement of Absorptivity and Emissivity 796
20.6.1. Spectral Measurements 796
20.6.2. Thermal Measurements 799
20.7. Spacecraft Thermal Balance 811
20.7.1. Deep Space Probes 811
20.7.2. Satellites 812
20.7.3. Thermal Design 821
20.8. Testing 823

20. Spacecraft Thermal Design

20.1. Definitions

The definitions given below are in accord with the notation of Worthing and Halliday [1]. This notation and the applied concepts have appeared in a recent report by some employees of NBS [2] and are rapidly being adopted throughout the aerospace industries, although they are not necessarily accepted by others. These are related parenthetically to the notation in Chapter 2.

20.1.1. Thermal Radiation Properties Terms [1-3].

Symbol	*Term*	*Units*	*Definition*
R, W	Radiancy (Emittance)	w cm^{-2}	The rate of radiant energy emission from a unit area of a source in all the radial directions of the overspreading hemisphere. The superscript b or 0 indicates a blackbody value.
N	Steradiancy (Radiance)	w cm^{-2}	The rate of radiant energy emission per unit solid angle per unit of projected area of a source, in a stated angular direction from the surface.
H	Irradiance	w cm^{-2}	The power per unit area incident on a surface.
ϵ	Emittance (Emissivity)	–	The ratio of the rate of radiant energy emission from a body, as a consequence of its temperature only, to the corresponding rate of emission from a blackbody at the same temperature.
α	Absorptance	–	The ratio of the radiant energy absorbed by a body to that incident upon it.
ρ	Reflectance	–	The ratio of the radiant energy reflected by a body to that incident upon it.
ϵ, α, ρ	Emissivity, Absorptivity, Reflectivity	–	Special cases of emittance, absorptance, and reflectance; these are fundamental properties of a material that has an optically smooth surface and is sufficiently thick to be opaque.
τ	Transmittance	–	The ratio of the radiant energy transmitted through a body to that incident upon it.

ϵ, α, ρ, and τ require additional qualification for precise definition. The following terms are used as modifiers, and the symbols are used as subscripts: λ = spectral; T = total; H = hemispherical; N = normal. In the list below these modifiers are combined with emittance as an example. Each term denotes surface temperature referenced to a blackbody at the same temperature.

ϵ_λ ..... Spectral emittance	Ratio of spectral radiancy (or monochromatic radiancy at a given wavelength) from a body to that of a blackbody.
ϵ_T ..... Total emittance	Total radiancy (radiation over the entire spectrum of emitted wavelength) from a body to that of a blackbody.
ϵ_H ..... Hemispherical emittance	Radiancy from a body to that of a blackbody.
ϵ_θ ..... Directional emittance	Steradiancy from a body to that of a blackbody at an angle θ.
ϵ_N ..... Normal emittance	The special case of directional emittance when the emittance is in a direction normal to the surface.

For more precision, subscripts are combined:

ϵ_{TH} ..... Total hemispherical emittance

ϵ_{TN} ..... Total normal emittance

$\epsilon_{\lambda H}$ ..... Spectral hemispherical emittance

$\epsilon_{\lambda N}$ ..... Spectral normal emittance

20.1.2. Space Technology Terms.

Albedo (Reflected Solar Radiation)	The ratio of the total solar radiant energy returned by a body to the total solar radiant energy incident on a body.
Insolation (Solar Radiation)	The irradiation of a body by direct total solar radiant energy.
Earth Shine (Earth Radiation)	The total radiancy from the earth as a consequence of its apparent temperature.
Celestial Sphere	A sphere of infinite radius whose center is the center of the earth, and upon which appear projected the stars and other astronomical bodies.
Ecliptic	The great circle on the celestial sphere formed by its intersection with the plane of the earth's orbit.
Ecliptic Plane	The plane defined by the orbit of the earth about the sun.
Planet Inclination	The angle between the orbit of the planet about the sun and the ecliptic plane.
Equatorial Inclination	The angle between the planet's equator and the ecliptic.

20.2. Basic Heat-Transfer Theory

20.2.1. Conduction. The general conduction equation holds in space as in a gravity field, although surface interface conductances may be altered due to the effects of the vacuum.

$$P = q_{1-2} = \frac{k}{x} A(T_1 - T_2) = hA(T_1 - T_2)$$

where	k = thermal conductivity	Btu (hr ft °F)$^{-1}$	w(cm °C)$^{-1}$
	x = thermal path	ft	(cm)
	A = cross-section area	ft^2	(cm^2)
	T = temperature	°F	(°K)
	$\frac{k}{x} = h$ = thermal conductance	Btu (hr ft^2 °F)$^{-1}$	w cm^{-2} (°C)$^{-1}$
	q, P = heat flow, power	Btu hr^{-1}	w

Thermal joint conduction in a vacuum has been analyzed by Fried [4] and by Fried and Costello [5], and some typical examples are presented in Fig. 20-1, 20-2, and 20-3.

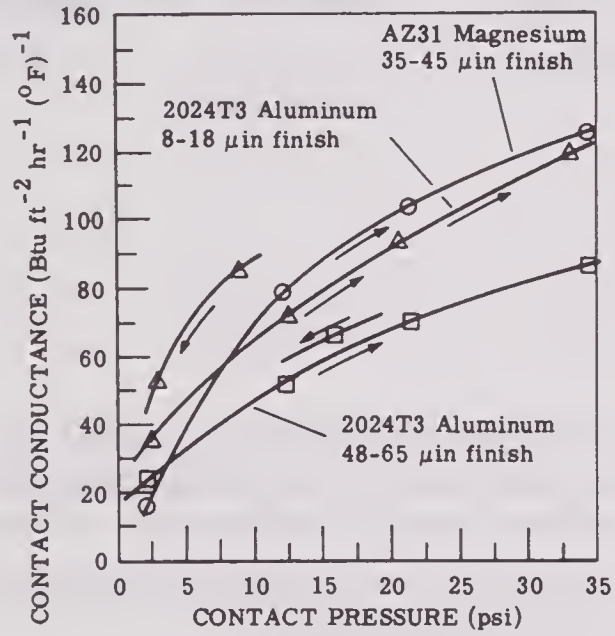


FIG. 20-1. Thermal contact conductance for aluminum and magnesium joints in vacuum [5].

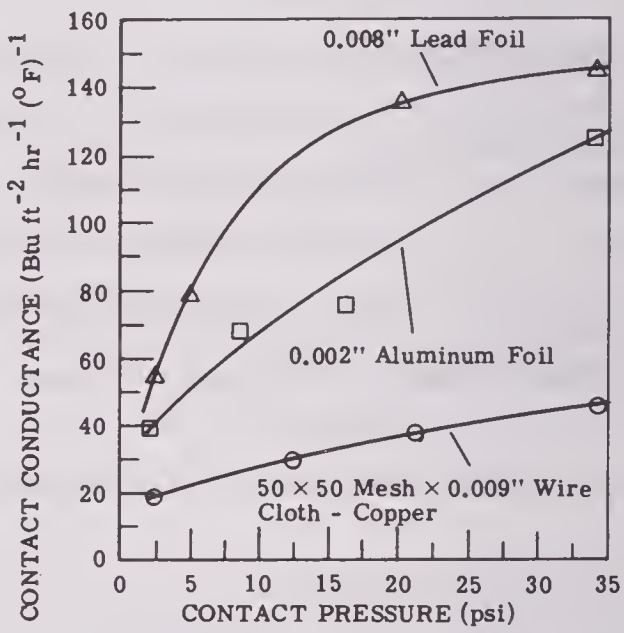


FIG. 20-2. Thermal contact conductance for aluminum joint with metallic shims in vacuum [5].

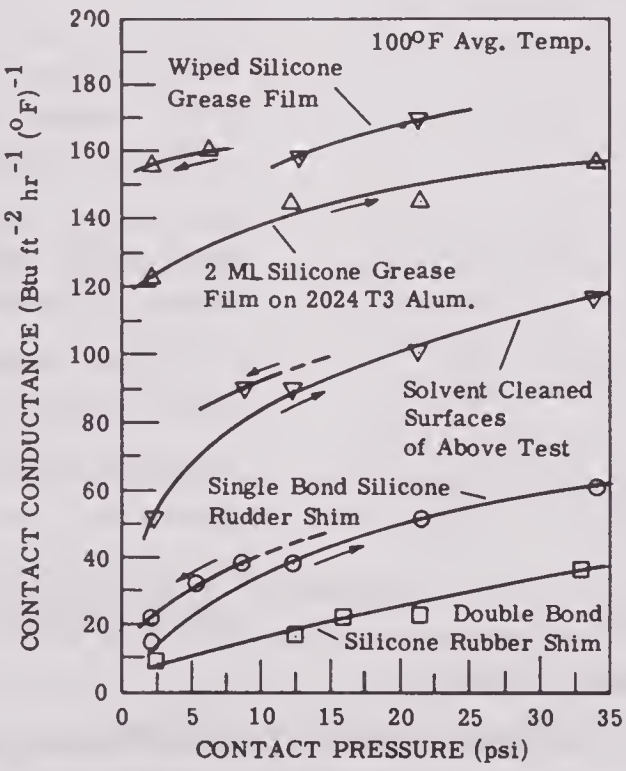


FIG. 20-3. Thermal contact conductance for nonmetallic shims in vacuum.

20.2.2. Radiant Heat Transfer Between Two Surfaces.

$$P_{1-2} = A_1 F_A F_\epsilon \sigma (T_1^4 - T_2^4)$$

where F_A and F_ϵ are as defined in Table 20-1 [6], σ is the Stefan-Boltzmann constant, and the other terms are as previously defined.

TABLE 20-1. VALUES OF F_A AND F_ϵ*

	F_A	F_ϵ
Surface A_1 small compared with the totally enclosing surface A_2	1	ϵ_1
Surfaces A_1 and A_2 of parallel discs squares, 2:1 rectangles, long rectangles	Fig. 20-4	$\epsilon_1\epsilon_2$
Surface dA_1 and parallel rectangular surface A_2 with one corner of rectangle above dA_1	Fig. 20-5	$\epsilon_1\epsilon_2$
Surfaces A_1 or A_2 of perpendicular rectangles having a common side	Fig. 20-6	$\epsilon_1\epsilon_2$
Surfaces A_1 and A_2 of infinite parallel planes, or surface A_1 of a completely enclosed body is small compared to with A_2	1	$\dfrac{1}{\left(\frac{1}{\epsilon_1}+\frac{1}{\epsilon_2}\right)-1}$
Concentric spheres or infinite concentric cylinders with surfaces A_1 and A_2	1	$\dfrac{1}{\frac{1}{\epsilon_1}+\frac{A_1}{A_2}\left(\frac{1}{\epsilon_2}-1\right)}$

*For more information see [7] and [8].

20.3. Thermal Coatings for Spacecraft

20.3.1. Optical Design. The thermal control of space vehicles depends ultimately on the radiation exchange of their surfaces with their environment. Adjustments depend largely upon the availability of materials with desirable spectral absorptance (emittance) characteristics. Materials which absorb little or no solar radiation and absorb (or emit) much radiation characteristic of bodies at about 300°K are very useful. In general, no material does this completely, but the degree can be specified by an α/ϵ ratio:

$$\alpha_s/\epsilon_{th} = \frac{\int_{\text{solar}} H_\lambda \alpha_\lambda \, d\lambda \Big/ \int_{\text{solar}} H_\lambda \, d\lambda}{\int_{300°K} R_{b\lambda} \epsilon_\lambda \, d\lambda \Big/ \int_{300°K} R_{b\lambda} \, d\lambda}$$

where the subscript s represents solar and th represents thermal.

About 95.3% of the sun's energy falls in the spectral range from 0.3 to 2.5 μ, 1.2% below 0.3 μ, and 3.4% above 2.5 μ. See Fig. 20-7 to 20-9 for solar and atmospheric

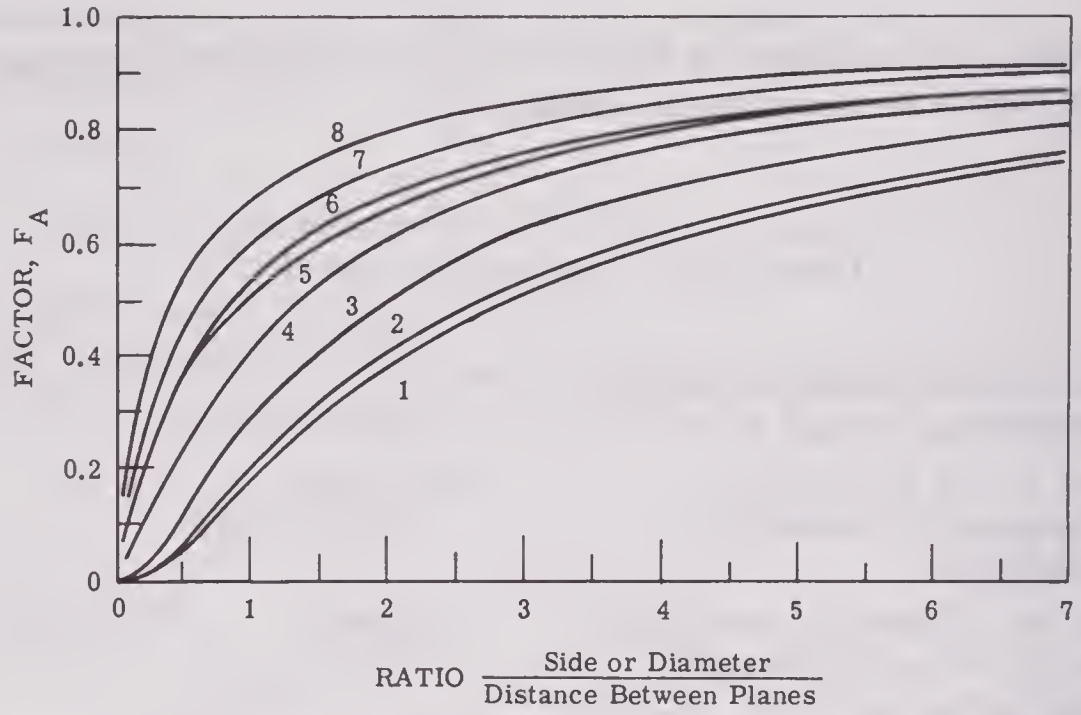


FIG. 20-4. Radiation between parallel planes, directly opposed. 1-2-3-4, direct radiation between the planes. 5-6-7-8, planes connected by non-conducting but reradiating walls. 1-5 are discs; 2-6 are squares; 3-7 are 2:1 rectangles; 4-8 are long narrow rectangles [9].

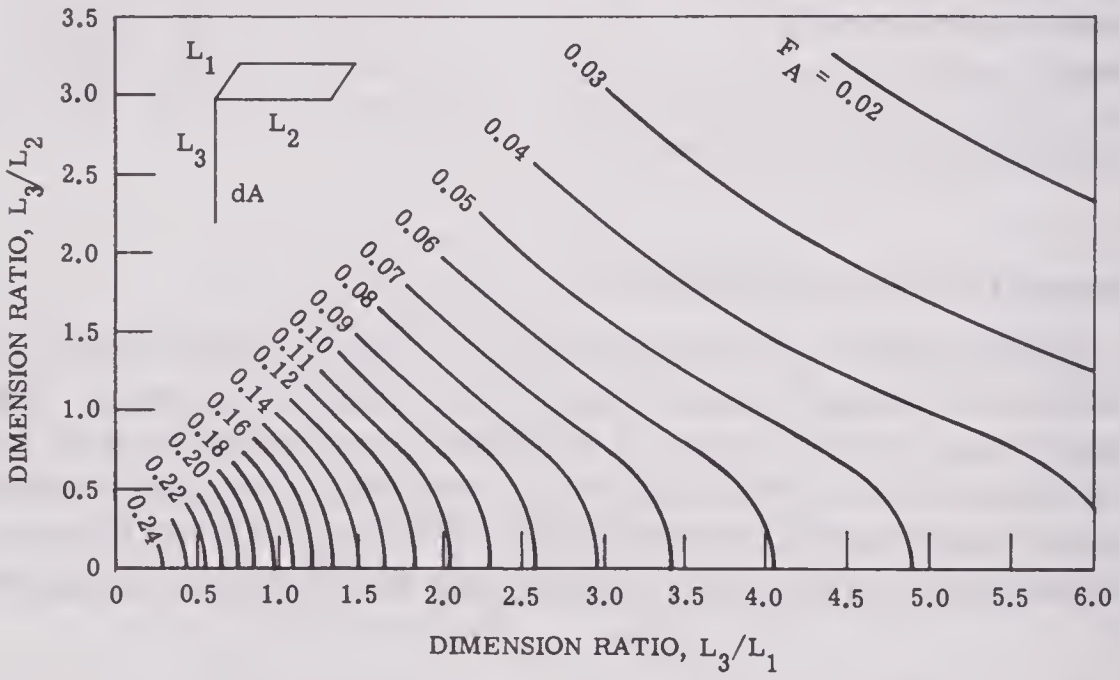


FIG. 20-5. Radiation between an element and a parallel plane. Radiation between surface element dA and rectangle above and parallel to it, with one corner of rectangle contained in normal to dA. L_1, L_2 = sides of rectangle; L_3 = distance from dA to rectangle; F_A = fraction of direct radiation from dA intercepted by rectangle [9].

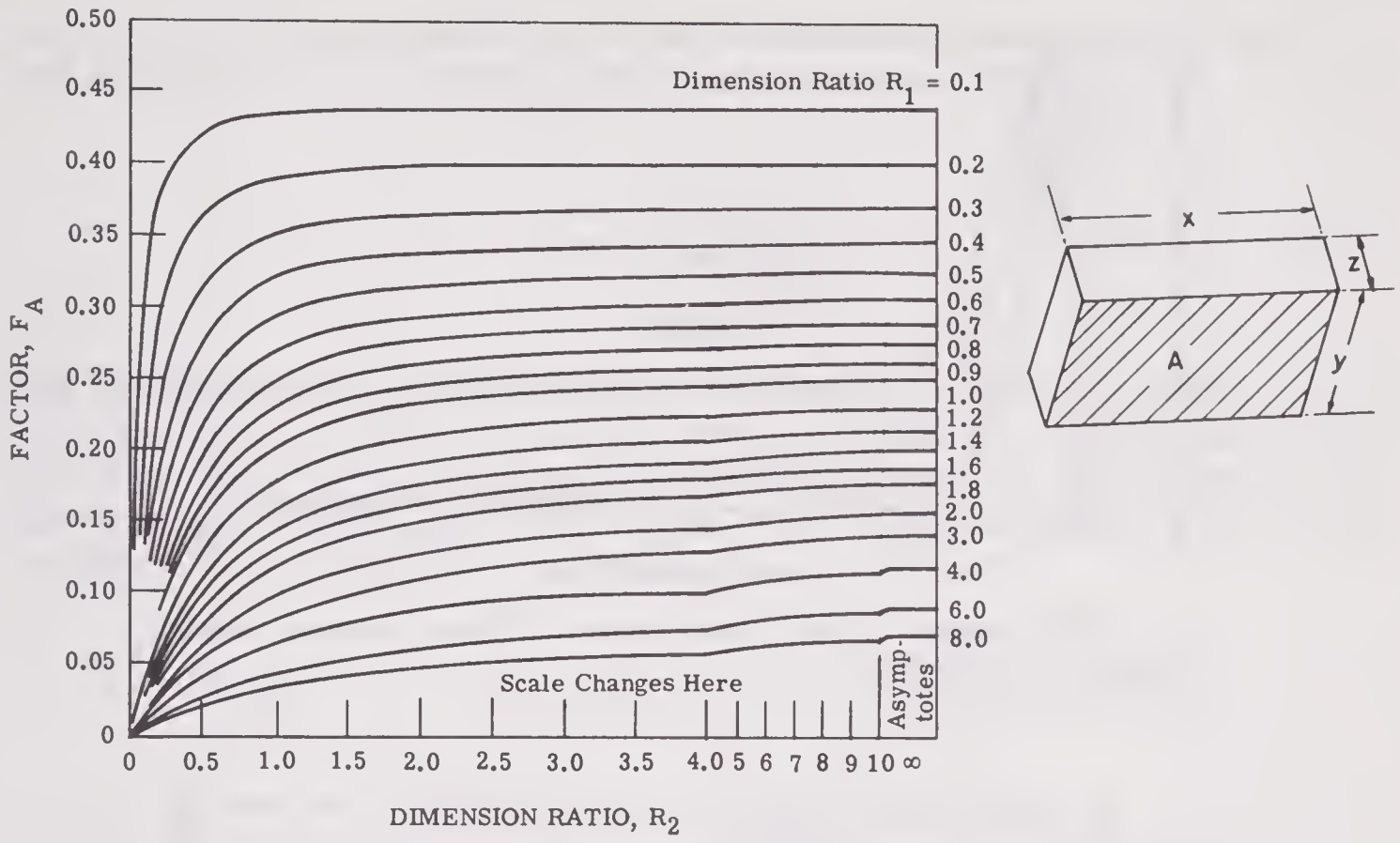


FIG. 20-6. Radiation between perpendicular planes. Radiation between adjacent rectangles in perpendicular planes. R_1 is the ratio of the length of the unique side of that rectangle on whose area the heat-transfer equation is based to the length of the common side, y/s in figure. R_2 is the ratio of the length of the unique side of the other rectangle to the length of the common side [9].

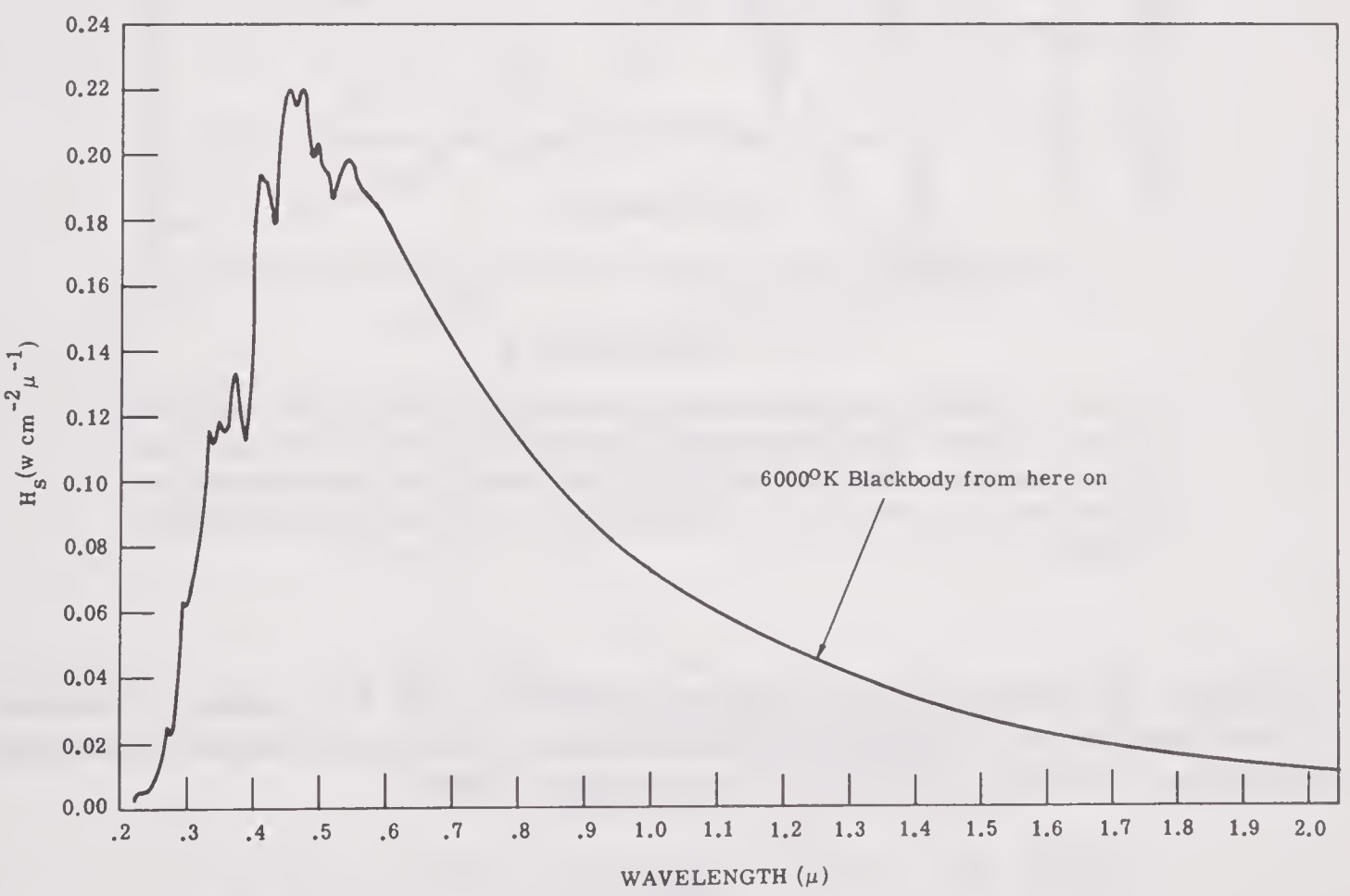


FIG. 20-7. The Johnson solar spectral curve [10].

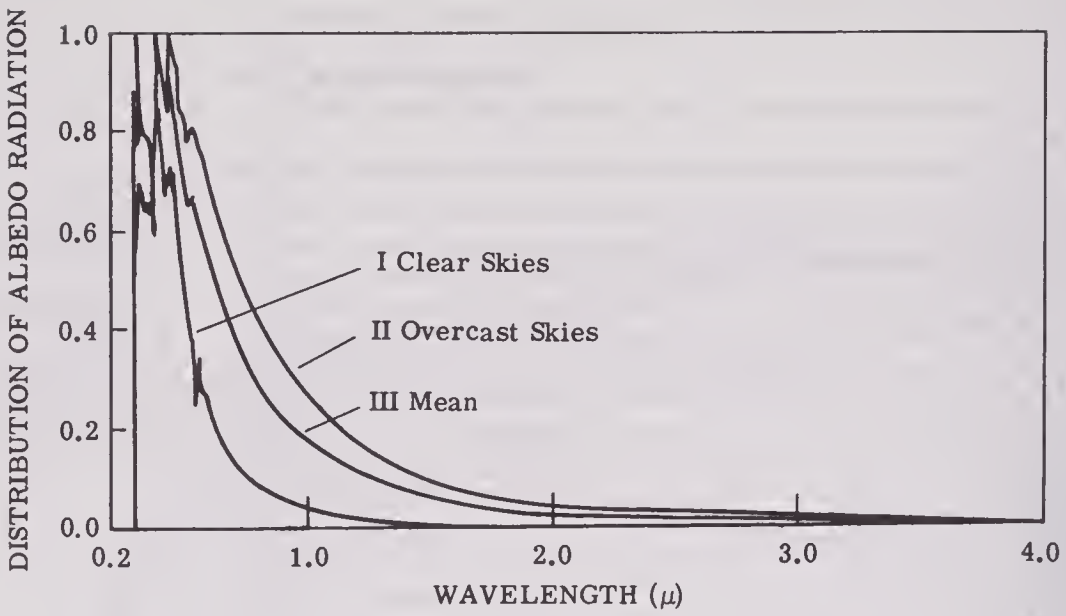


FIG. 20-8. Relative spectral distributions of albedo radiation under various sky conditions.

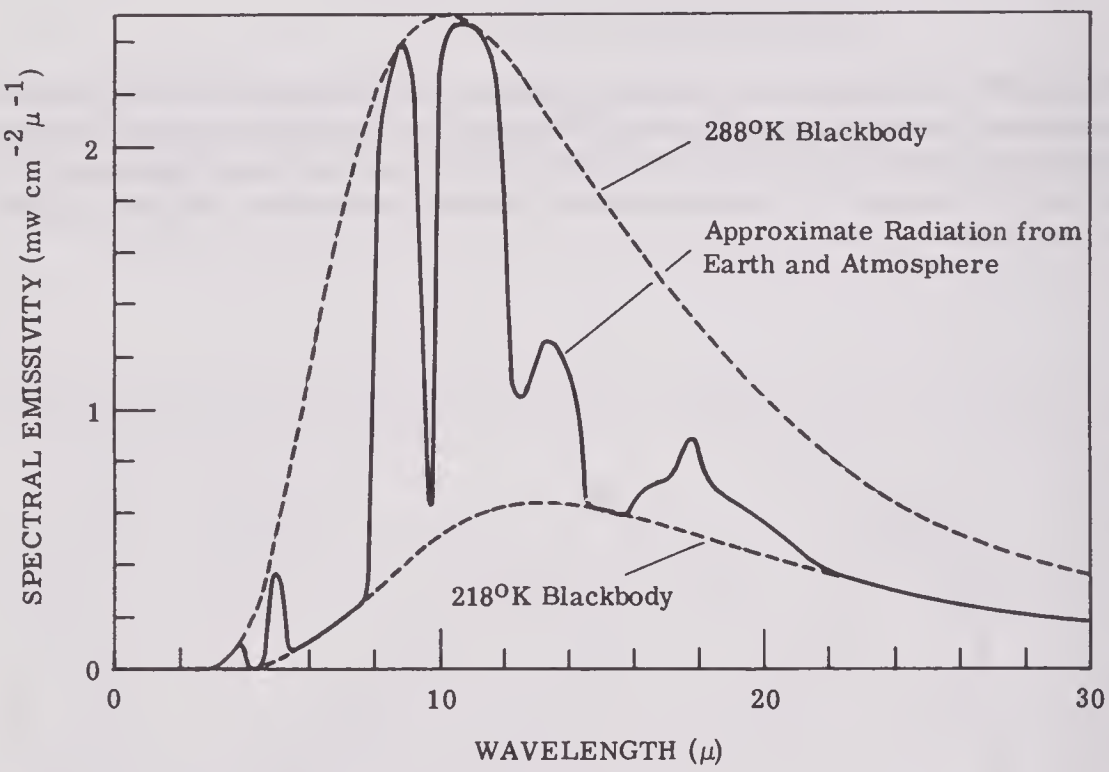


FIG. 20-9. A typical spectral emissive power curve for the thermal radiation leaving the earth. (The 288°K blackbody curve approximates the radiation from the earth's surface, and the 218°K blackbody curve approximates the radiation from the atmosphere in those spectral regions where the atmosphere is opaque.)

radiations. The maximum spectral energy is emitted in the 0.5-μ region. Therefore, the solar absorptance α_s of a coating depends upon the thermal radiation properties of the coating in the 0.3 – 2.5-μ range of the spectrum. Thus

$$\alpha_s = \frac{\int_{0.3\mu}^{2.5\mu} \alpha_\lambda H_\lambda \, d\lambda}{\int_{0.3\,\mu}^{2.5\,\mu} H_\lambda \, d\lambda} \tag{20-1}$$

About 0.001% of the radiation from a 300°K blackbody falls below 2.5 μ and about 95% falls below 40 μ. Therefore, a definition of emittance at 300°K is*

$$\epsilon_{th} = \frac{\int_{2.5\mu}^{40\mu} \epsilon_\lambda R_{b\lambda}\, d\lambda}{\int_{2.5\mu}^{40\mu} R_{b\lambda}\, d\lambda} \tag{20-2}$$

The 40-μ upper limit is arbitrary and can be altered as satellites with lower temperatures and longer wavelength blackbody curves are considered. The emissivities of some materials which have high and low ratios, and some which are also useful as flat absorbers and reflectors, are shown in Fig. 20-10. Characteristics of optimum and practical surfaces are shown in Fig. 20-11. See also Chapter 8.

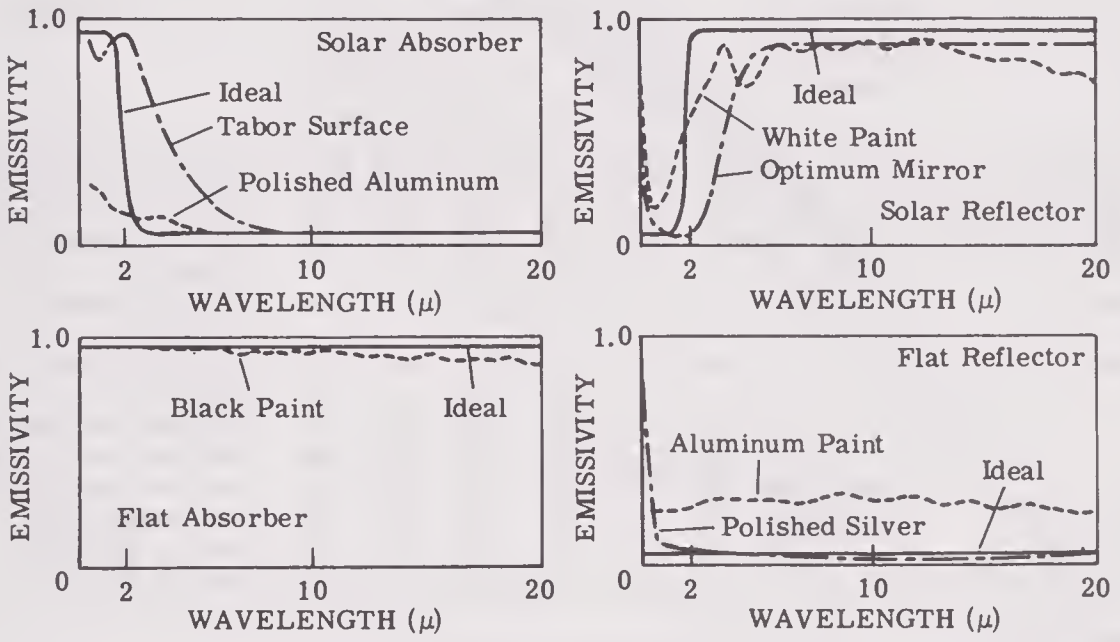


FIG. 20-10. Representative spectral emissivity curves for four ideal surfaces.

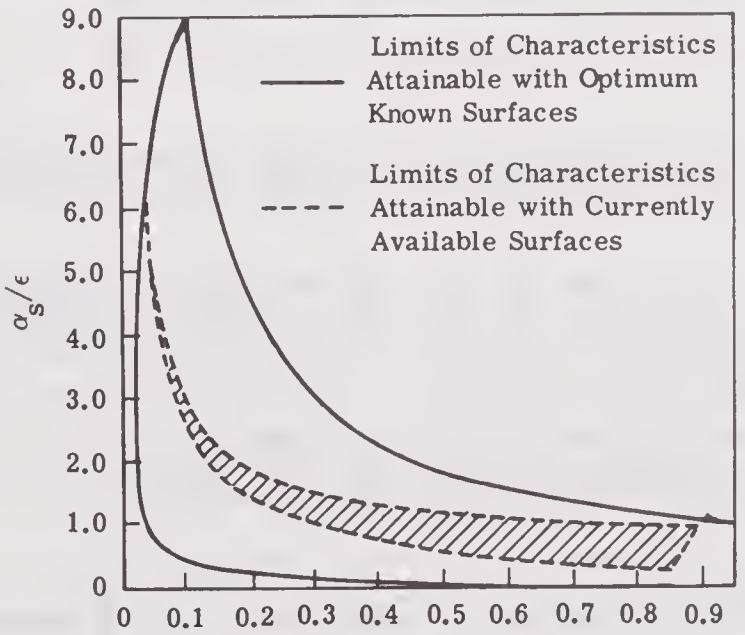


FIG. 20-11. Radiation characteristics attainable with optimum surfaces compared to those attainable with currently available materials.

*The ratio will often be written simply as α/ϵ rather than α_s/ϵ_{th}.

20.3.2. Environmental Consideration. Aside from the theoretical analysis of thermal control systems, particular emphasis must be placed on the practical aspects involved in manufacture, storage, and launch conditions that have adverse effects on the assumed values of α_s/ϵ_{th}.

Extreme care must be taken to prevent scoring, fingerprints, oil films, etc., on the thermal coating; these will alter the α_s/ϵ_{th} ratio as prescribed by analyses.

Ultraviolet radiation degrades the transmission of some films. A series of ultraviolet degradation values is presented in Table 20-2 [8].

Little work has been done on the degrading effects that meteoroid bombardment has on coatings, although data on the particle environment are available (Fig. 20-12).

TABLE 20-2. ULTRAVIOLET DEGRADATION OF α_s

Material Description	*Exposure*[a]			*Optical Data*[b]						*Remarks*
	I	*t*	*E*	*Initial*			*Final*			
White Paints				α_s	ϵ	α_s/ϵ	α_s	ϵ	α_s/ϵ	
Sicon 7 × 1153	6	12	72	0.25	0.89	0.283	0.33	0.88	0.38	4 coats on Dow 15 on Mg HM-21A
Sicon 7 × 1153	6	100	600	0.26	0.90	0.29	0.37	0.82	0.45	3 coats on Dow 15 on Mg HM-21A
Sicon 7 × 1153	6	100	600	0.31	0.89	0.35	0.59	0.88	0.67	4 coats on Dow 17 on Mg HM-21A
Skyspar A-423	6	25	150	0.23	0.89	0.26	0.36	0.91	0.4	4 coats on P-323-B on Dow 17 on Mg HM-21A
Skyspar A-423	6	46	276	0.26	0.93	0.28	0.37	0.92	0.4	3 coats on P-323-B on Dow 17 on Mg HM-21A
Skyspar A-423	6	12	72	0.28	0.9	0.31	0.31	0.90	0.34	5 coats on P-323-B on Dow 17 on Mg HM-21A
Kemacryl	6	25	150	0.27	0.80	0.34	0.32	0.80	0.4	3 coats on Dow 17 on Mg HM-21A
M49WC 17	6	25	150	0.26	0.79	0.33	0.32	0.80	0.41	4 coats on Dow 17 on Mg HM-21A
M49WC 17	6	100	600	0.26	0.79	0.33	0.33	0.82	0.40	4 coats on Dow 17 on Mg HM-21A
Fuller 517-W-1 Silicone	6	25	150	0.33	0.82	0.40	0.35	0.84	0.42	2 mils on Mg HM-21A
Fuller 517-W-1 Silicone	6	25	150	0.28	0.87	0.31	0.30	0.85	0.36	4 mils on Mg HM-21A
Fuller 517-W-1 Silicone	6	46	276	0.33	0.85	—	0.29	0.87	0.34	3 coats on Mg HM-21A
LMSC Research Paints										
Sodium Silicate "D" + Ultrox	6	91	546	0.29	0.82	0.34	0.32	0.89	0.37	
Sodium Silicate "D" + Ultrox	6	119	714	0.29	0.91	0.31	0.29	0.85	0.33	
Sodium Silicate "D" + Ultrox	6	47	282	0.27	0.87	0.31	0.47	0.86	0.55	
Black Paints										
Kemacryl M49 BC12	6	100	600	0.94	0.88	1.07	0.92	0.84	1.10	4 coats on Dow 17 on Mg HM-21A
Micobond	6	105	630	0.94	0.91	1.04	0.98	0.87	1.13	4 coats on Dow 17 on Mg HM-21A
Al Silicone, 10043										
Fuller 171-A-152	6	100	600	0.22	0.16	1.34	0.33	0.24	1.35	2 mils
Fuller 171-A-152	6	100	600	0.21	0.16	1.26	0.30	0.20	1.48	4 mils
Anodized & Misc. Finishes										
Dow 15 on Mg HM-21A	6	12	72	0.23	0.05	4.47	0.28	0.05	5.22	
Dow 15 on Mg HM-21A	6	105	630	0.17	0.07	2.54	0.37	0.34	1.09	
Dow 15 on Mg HM-21A	20	12	240	0.18	0.08	2.41	0.38	0.11	3.39	
Rokide A (10546) on aluminum	6	95	570	0.31	0.75	0.40	0.44	0.74	0.59	
Evaporated gold on 1/16" Kel-F	6	119	714	0.30	0.07	4.05	0.31	0.09	3.45	Severe discoloration of gold with surface blistering of plastic.

[a] Exposure, *E*, (sun hours) is a product of intensity, *I*, (suns) and time, *t*, (hours).

[b] Optical data include solar absorptance, infrared total hemispherical emittance, and post-exposure data.

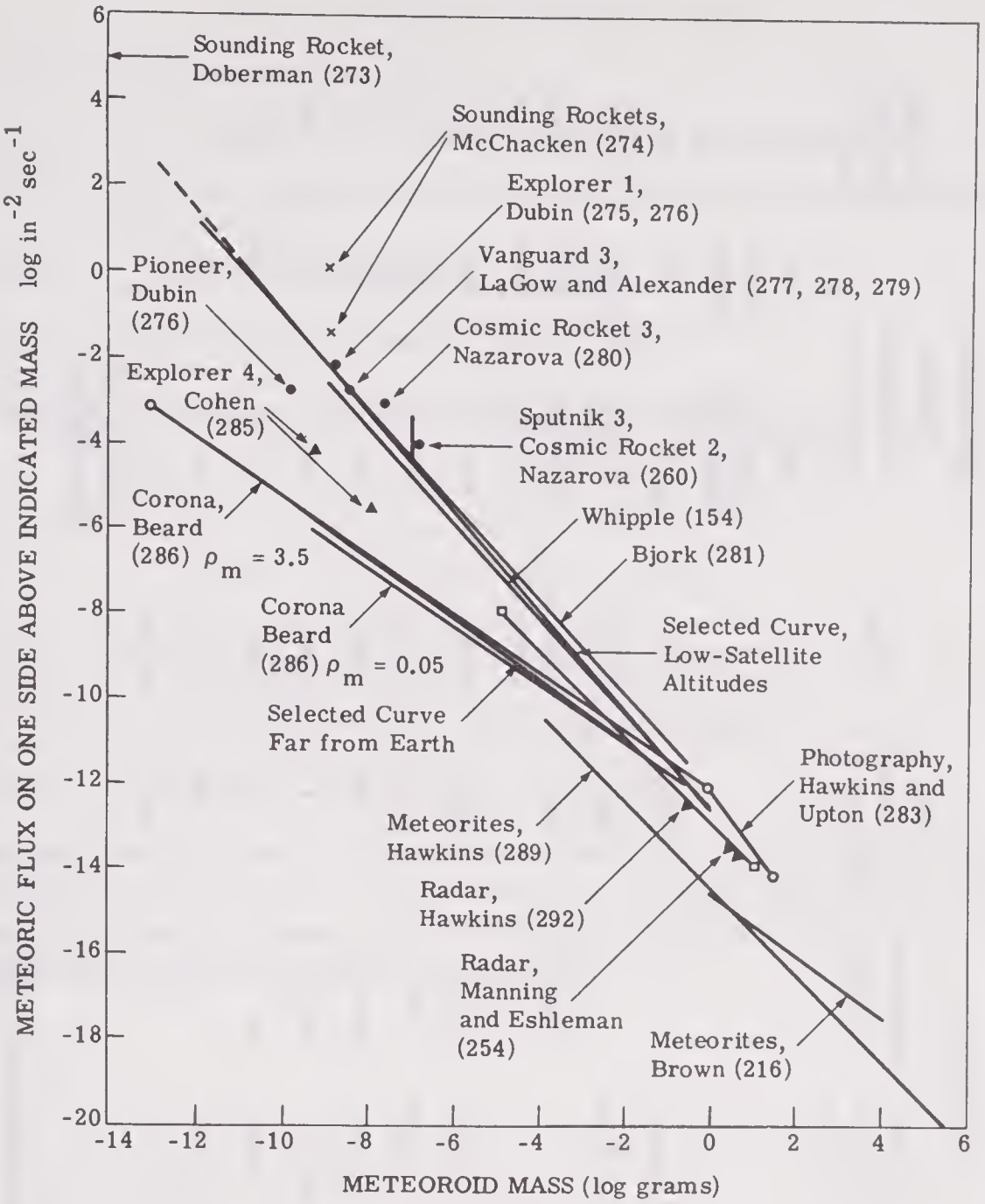


FIG. 20-12. Meteoric flux [11]. References on the graph are those of the original article.

20.4. Solar System

Table 20-3 contains astronomical constants which are useful in the thermal design of spacecraft.

20.5. Angular Dependence of Reflectivity

Total hemispherical emissivity and total normal emissivity for opaque surfaces depend upon how total directional reflectivity varies with the angle of incidence. The following set of equations (azimuthal symmetry case) describes the relationship of total hemispherical emissivity to the spectral directional reflectivity of a coating [18].

$$\epsilon_{TH}(T) \frac{\int_{\lambda_1}^{\lambda_2} \int_{\theta=0}^{\theta=\pi/2} \epsilon_\lambda(\theta) 2\pi \sin\theta \cos\theta \, d\theta \, W_\lambda^0(T) \, d\lambda}{\int_{\lambda_1}^{\lambda_2} \int_{\theta=0}^{\theta=\pi/2} 2\pi \sin\theta \cos\theta \, d\theta \, W_\lambda^0(T) \, d\lambda} \tag{20-3}$$

The total emittance of a nonblackbody at any given temperature is equal to its total absorptance of radiation from a blackbody at the same temperature; or

TABLE 20-3. ASTRONOMICAL CONSTANTS

Body	*Major Semi-Axis* (AU)[1] [12]	*Perihelion* (AU)[1] [12]	*Aphilion* (AU)[1] [12]	*Radius* (km) [12]	*Inclination*		*Solar Irradiance* (Btu hr^{-1} ft^{-2}) [13]	*Albedo*	*Black Sphere Surface Temperature*		*Planet Irradiance* (Btu hr^{-1} ft^{2})
					Orbital to Ecliptic (°:′:″)	*Equatorial to Ecliptic* (°:′:″)			°K [12]	°R [12]	
Sun				696,000					5800[3] [14] 6000[4] [14]	10,440 10,800	
Mercury	0.3871	0.3075	0.4667	2,500	7:0:14		2953.9	0.07	450	810	745
Venus	0.7233	0.7184	0.7282	6,200	3:23:39		846.3	0.76 [11]	700 [15]	1,260	4360
Earth	1.000	0.9833	1.0167	6,378.39		23:26:59	442.4	0.25 [14]	260 [16]	468	83.0
								0.36[2] [15]	250[2] [17]	450	70.9
Mars	1.524	1.381	1.666	3,310	1:50:51	25:12	190.5	0.15	225	405	46
Jupiter	5.203	4.951	5.455	69,880	1:18:20	3:7	16.4	0.51	122	220	4
Saturn	9.539	9.008	10.070	57,550	2:29:42	26:45	4.9	0.50	90	162	1.2
Uranus	19.18	18.28	20.09	25,500	46:23	97:59	1.2	0.66	64	115	0.3
Neptune	30.06	29.80	30.32	25,000	1:46:27	29:	0.5	0.62	51	92	0.1
Pluto	39.52	29.69	49.34	1,738	17:8:38		0.3	0.16	45	81	
Moon	384,400 km						442.2	0.124	243 Hot side: 374 Cold side: 120		

[1]1 astronomical unit (AU) = 149.4×10^6 km. The temperature of space is assumed to be 4°K.
[2]Grumman design for Orbiting Astronomical Observatory.
[3]Total emitted radiation.
[4]Visible and infrared radiation; ultraviolet not included.

$$\epsilon_{TH}(T) = \alpha_{TH}(T) \quad \text{(Kirchhoff's law)} \tag{20-4}$$

where $\epsilon_\lambda(\theta)$ = specular spectral emissivity

$\epsilon_{TH}(T)$ = total hemispherical at temperature T in degrees absolute

$\alpha_{TH}(T)$ = total hemispherical absorptivity for the same source temperature T in degrees absolute

$W_\lambda{}^0(T) = R_{b\lambda}(T)$ = Planck's blackbody function

λ_1, λ_2 = wavelength limits to cover 95% of $R_b(T)$

The equation for total normal emissivity which is readily measured in the laboratory, would be

$$\epsilon_{TN}(T) = \frac{\int_{\lambda_1}^{\lambda_2} \epsilon_\lambda(\theta) W_\lambda{}^0(T)\, d\lambda}{\int_{\lambda_1}^{\lambda_2} W_\lambda{}^0(T)\, d\lambda} \quad \theta = 0^0 \tag{20-5}$$

$$\epsilon_{TN}(T) = 1 - \rho_{TN}(T) \quad \text{(for opaque surface)} \tag{20-6}$$

$$\epsilon_{TN}(T) = \alpha_{TN}(T) \quad \text{(Kirchhoff's law)} \tag{20-7}$$

For metal surfaces the ratio $\epsilon_{TH}/\epsilon_{TN}$ is greater than 1 and for dielectric surfaces may vary from 0.92 to slightly above 1. The relation is given in Fig. 20-13 as a function of the measurable normal emissivity [8]. This is an empirical relation. For an ideal diffuse emitter (according to Lambert's law), $\epsilon_{TH}/\epsilon_{TN} = 1$.

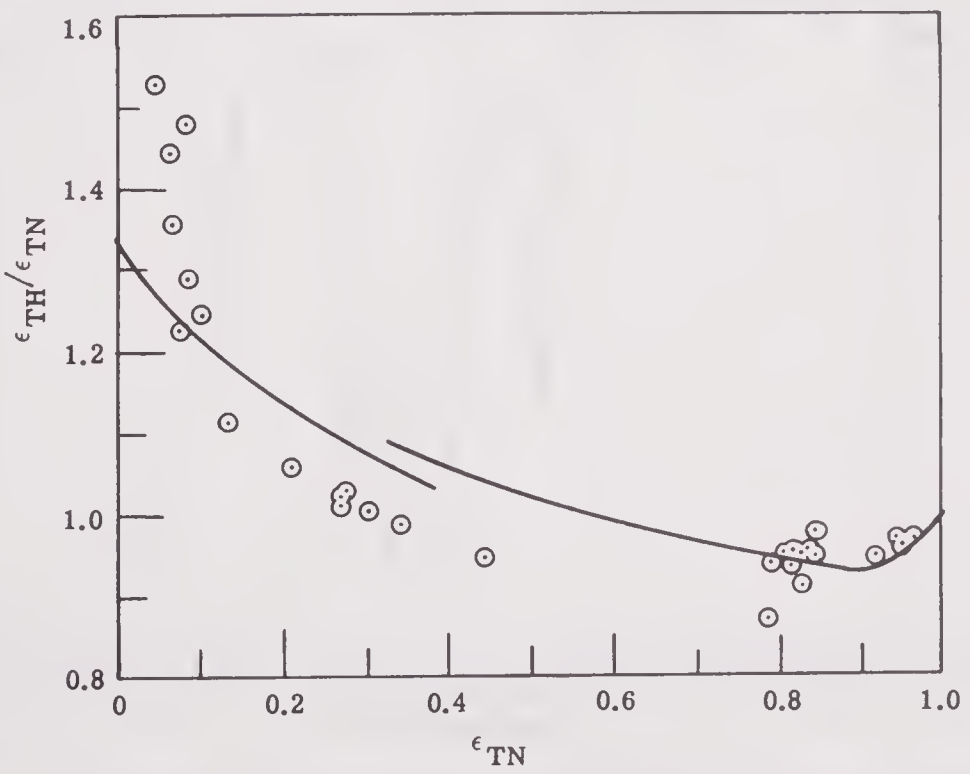


FIG. 20-13. Ratio of total hemispherical emissivity to normal emissivity. The curve is from Jakob [8].

20.6. Measurement of Absorptivity and Emissivity

Two distinct measurement techniques are employed for obtaining α_s and ϵ_{TH}, the spectral reflection technique and the thermal techniques.

20.6.1. Spectral Measurements. The spectral reflection method is based on the definition of reflectivity which requires that the reflected energy be measured over the entire overspreading hemisphere. In determining α_s, the definition is complied with in the 0.25-2.5-μ range using an integrating sphere-spectrophotometric apparatus [19]. The resultant property measured is the directional spectral reflectivity. Integration of the spectral reflectivity data over the Johnson curve (by IBM programming or equal-energy increment-summation techniques) yields total normal solar reflectivity. When this solar reflectivity is subtracted from one, α_s is realized.

On the basis of the reflectivity-reciprocity relation, the normal or directional spectral reflectivity in the 2.5-50-μ range may be measured with a heated-cavity reflectometer-spectrophotometer type of apparatus [20]. The sample is irradiated hemispherically and viewed at a specific angle by the spectrophotometer's optical system. Integration of the spectral reflectivity data over the appropriate blackbody curve characteristics of the spacecraft temperature gives the total normal emissivity properties of the material. The total normal emissivity must be converted to total hemispherical emissivity (use Fig. 20-12). For precise thermal-control design, the normal spectral reflectivity data should be supplemented with directional reflectivity information and an integration (or discrete summation) performed. The normal spectral reflectivity data for several materials are shown in Fig. 20-14 to 20-19 [21].

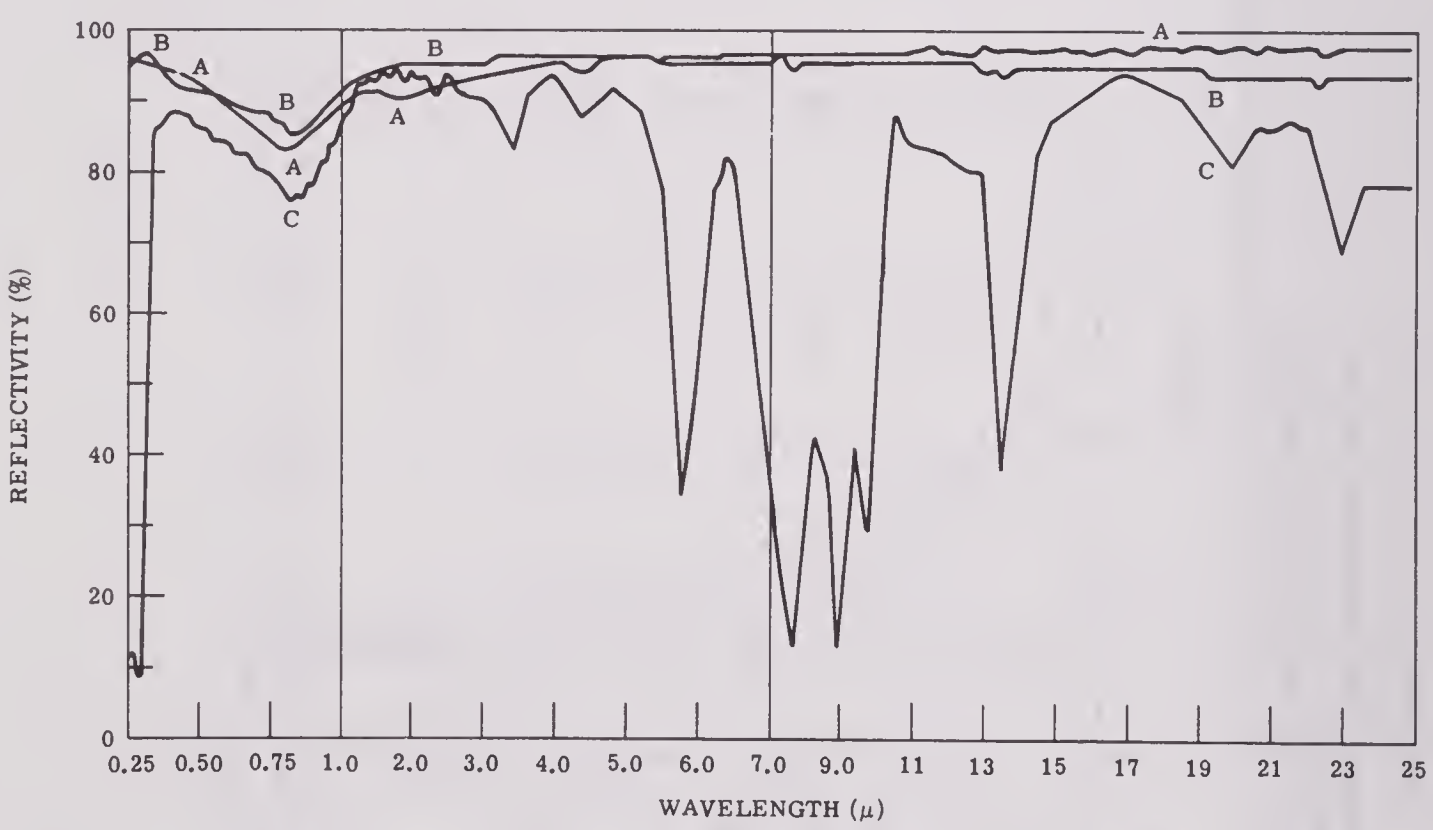


FIG. 20-14. Compiled spectral reflectivities. A = evaporated aluminum, 25 μin. on polished 6061-T6 aluminum. B = evaporated aluminum (0.2 μ), on 1/4-mil Mylar crumpled and stretched, (looking at aluminum). C = evaporated aluminum (0.2 μ) on 1/4-mil Mylar crumpled and stretched (looking at Mylar).

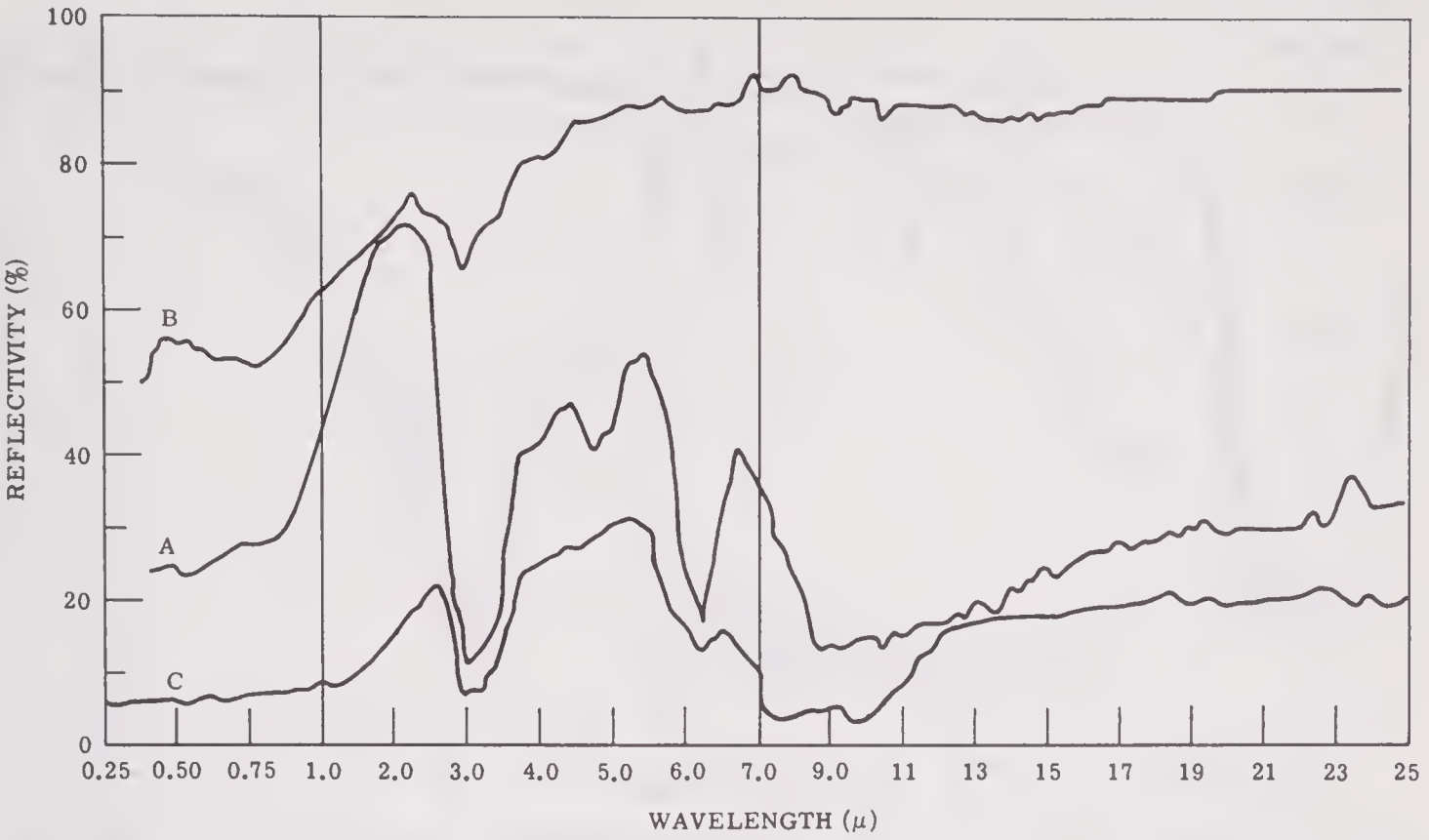


FIG. 20-15. Compiled spectral reflectivities. A = chromic acid anodize on 24S-T81 aluminum. B = sulfuric acid anodize on 24S-T81 extruded aluminum, chemically milled. C = hard anodize (1 mil) on 6061-T6 aluminum (35 amp ft^{-2}, 45 volts, 20°F).

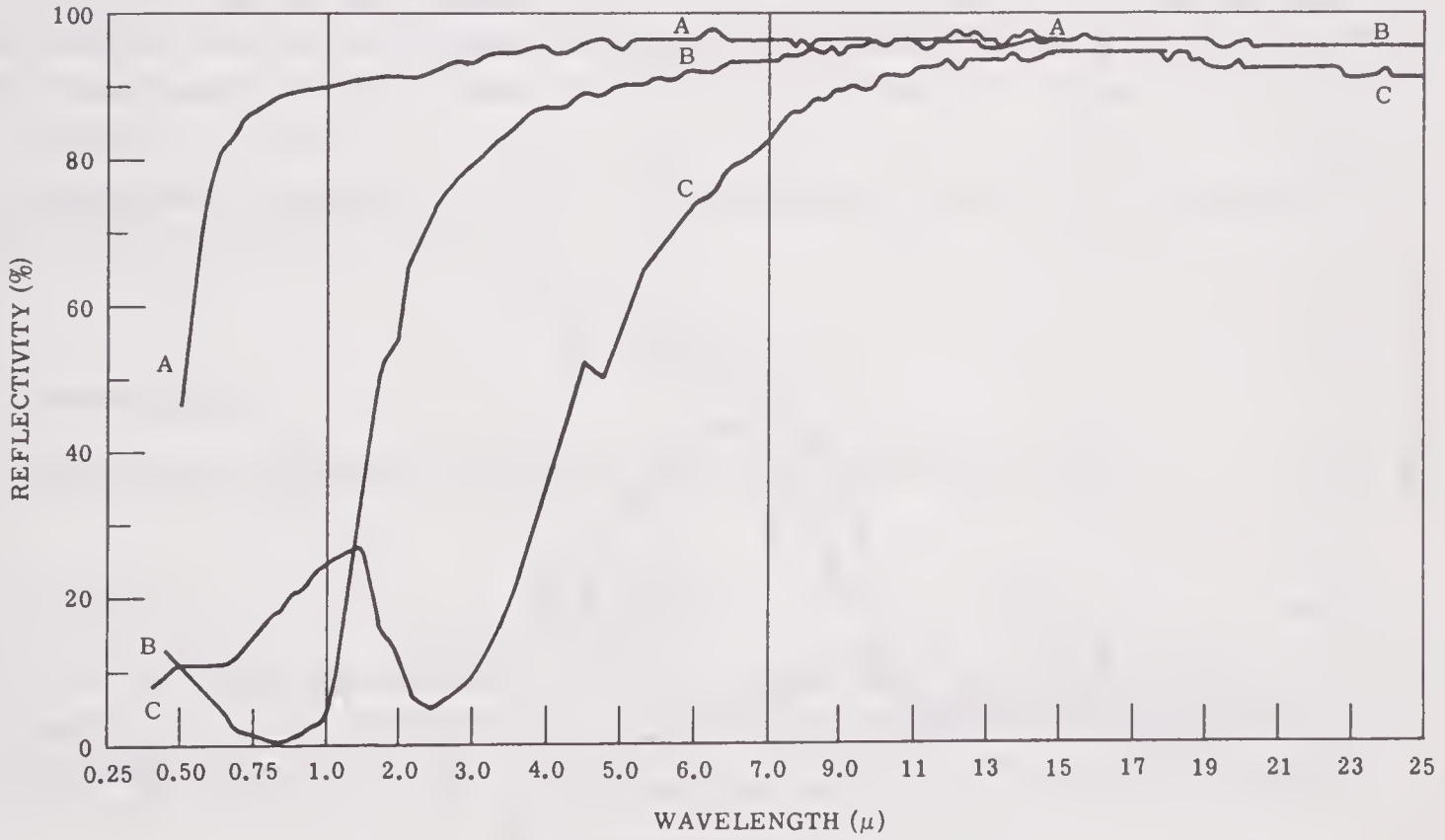


FIG. 20-16. Compiled spectral reflectivities. A = polished copper, 17 mils thick. B = Tabor solar collector chemical treatment (110-30) on nickel-plated copper. C = Tabor solar collector chemical treatment (125-30) on nickel-plated copper.

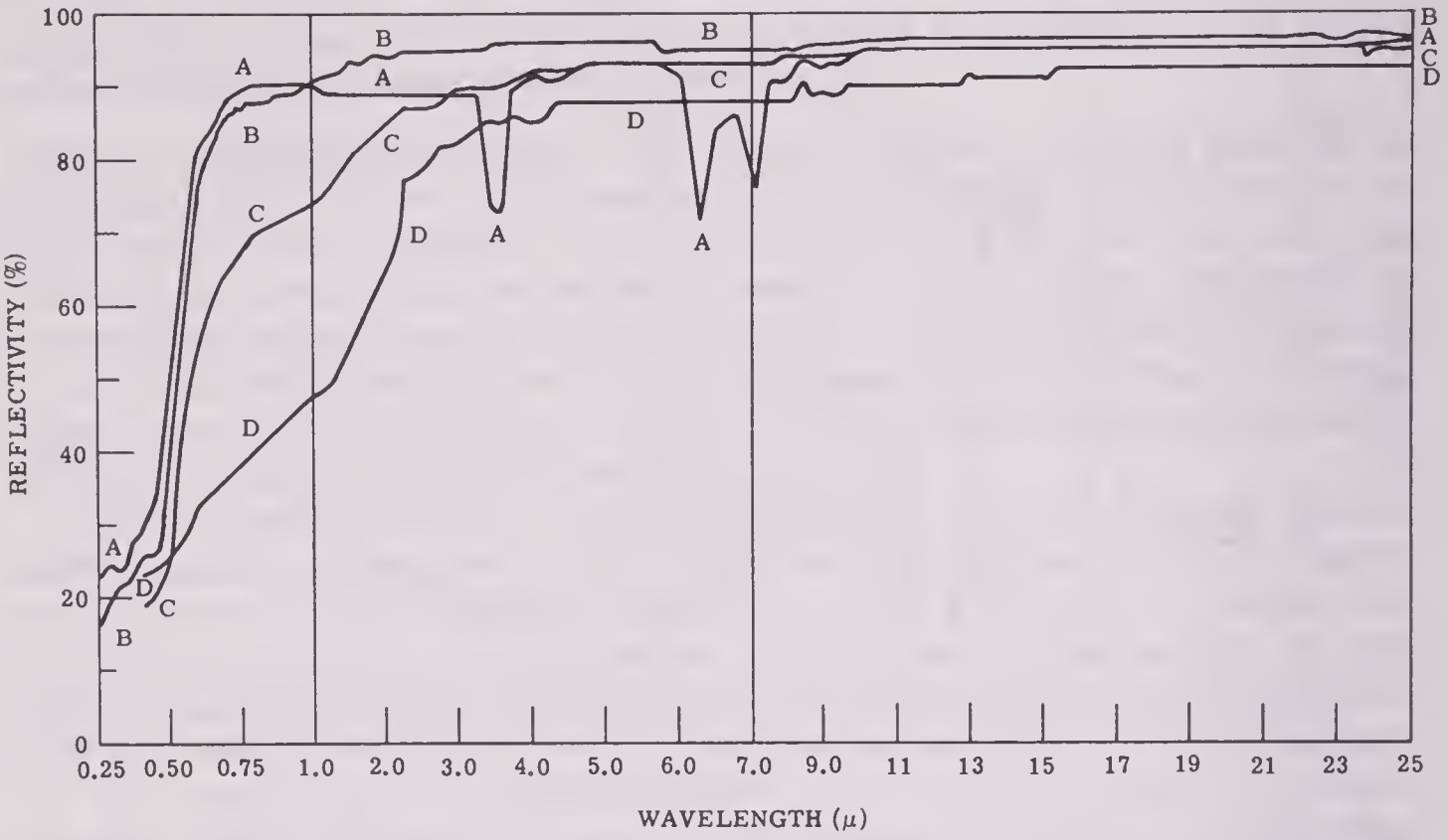


FIG. 20-17. Compiled spectral reflectivities. *A* = immersion gold, approximately 0.03 mil, on nickel-plated copper plate on 6061-T6 polished aluminum, aged 6 months in air, unpolished. *B* = vacuum-evaporated gold on fiberglass laminate. *C* = gold ash (80 μ) on 0.4-mil silver on Epon glass. *D* = white gold on polished MIL-S-5059 steel.

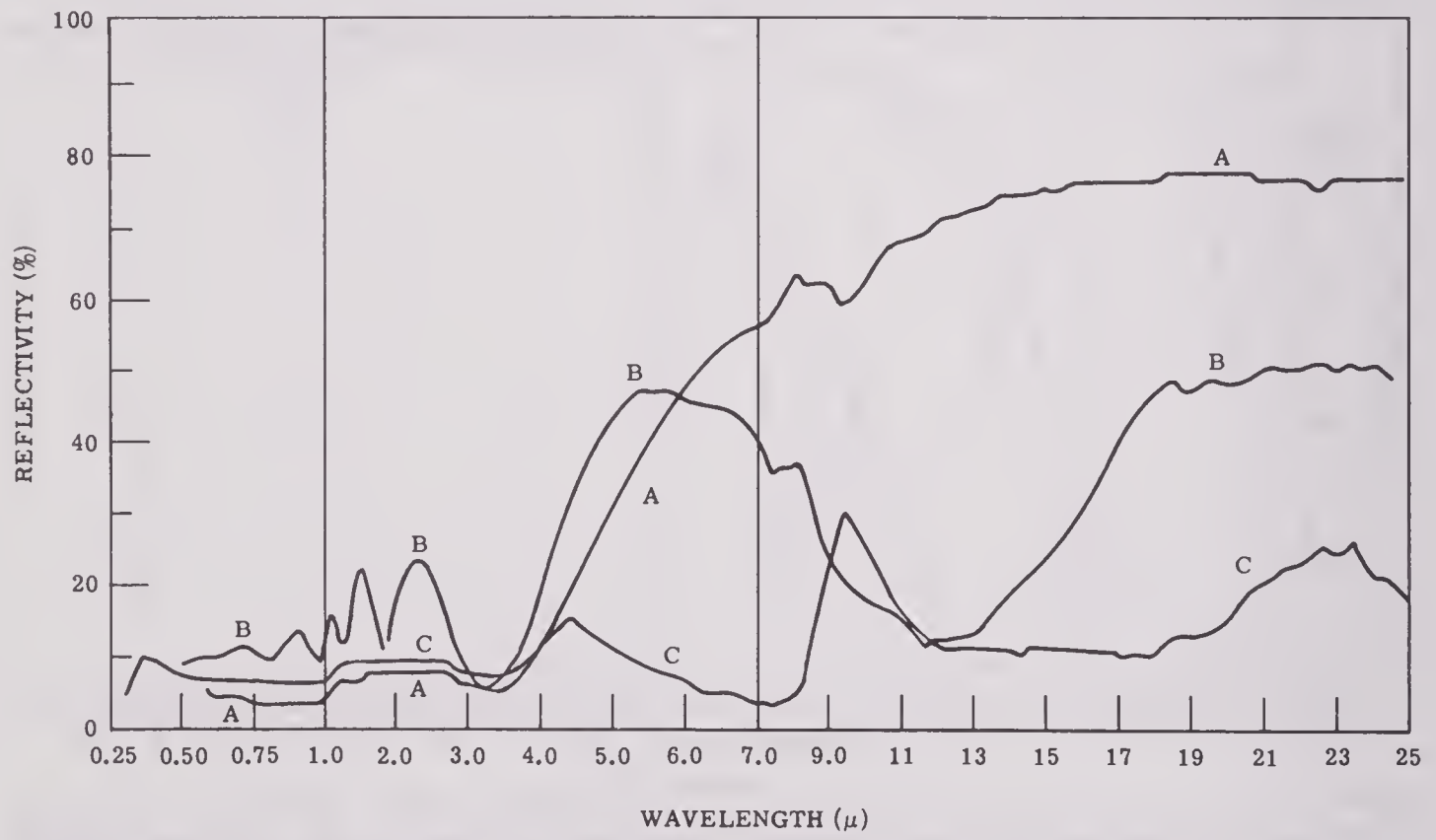


FIG. 20-18. Compiled spectral reflectivities. *A* = silicon solar cell, International Rectifier Corp. *B* = silicon solar cell, International Rectifier Corp. 1.11-μ evaporated coat, SiO—fast deposition rate. *C* = silicon solar cell, Hoffman Corp. Type 120-C, 3-mil glass cover.

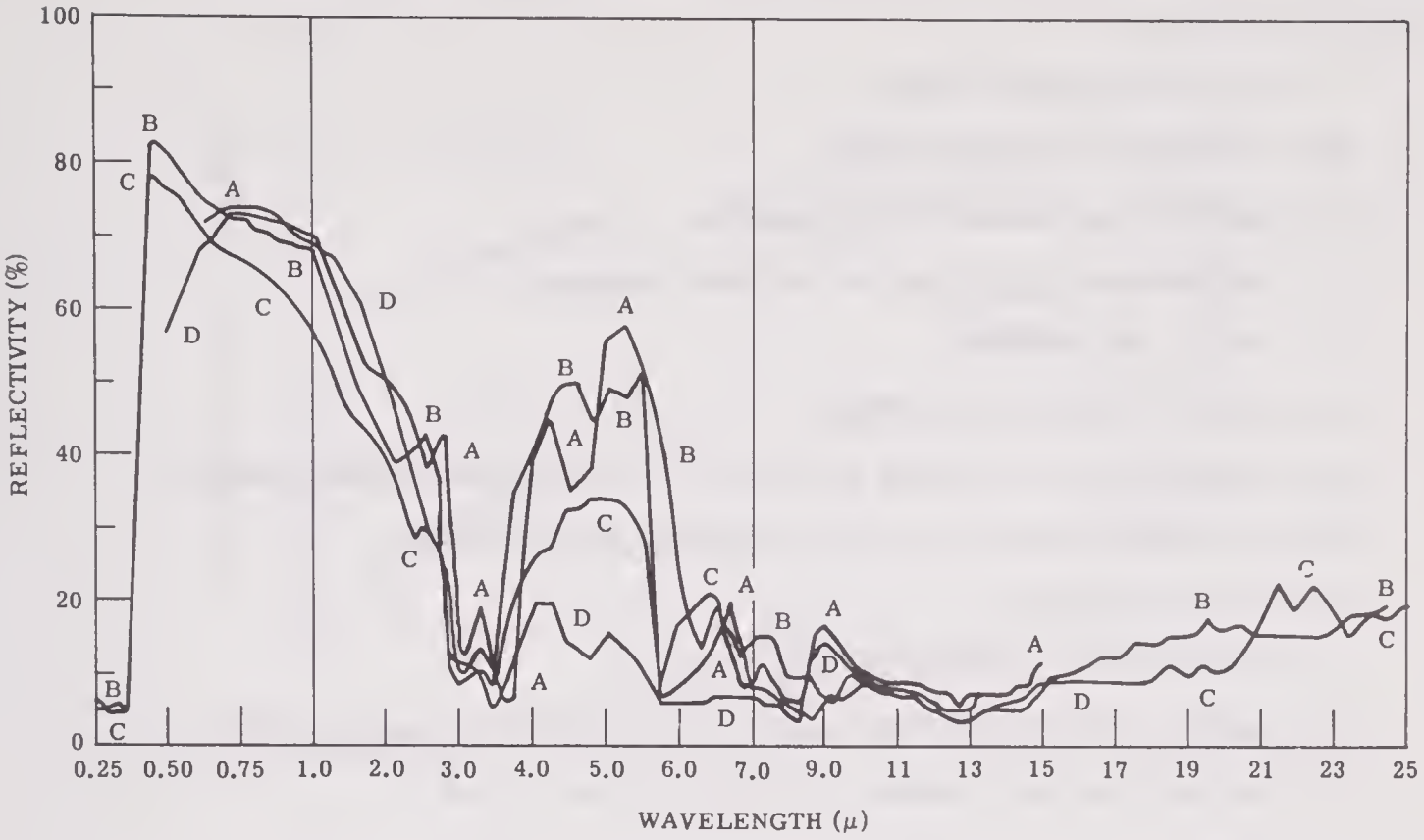


FIG. 20-19. Compiled spectral reflectivities. A = Flat white paint, Fuller No. 2882 on 2-mil polished aluminum. B = white epoxy resin paint of Mg. C = flat white acrylic resin, Sherwin-Williams M49WC8-CA-10144; MIL-C-15328A pretreatment wash coating on 22-mil 301 stainless steel 1/2 hard. D = white paint mixed with powdered glass, 7 mil on polished aluminum.

20.6.2. Thermal Measurements. Thermal measurement of α_s and ϵ_{th} will yield ϵ_{TH} directly, but the α_s measurement's validity depends on matching the solar simulator source with the spectral distribution of the correct spectrum. A normalized spectral curve of the solar carbon arc and a bare xenon lamp shows that the carbon arc simulates the solar spectrum more closely than other light sources. Measurements can be equilibrium or dynamic.

Equilibrium Measurements [18, 21]. With electrical heating and the property ϵ_{TN}, the governing relation is

$$\epsilon_{TH} = \frac{I^2 R}{\sigma(T_2^4 - T_1^4)}$$

With a solar simulator as the heater and the property α_s/ϵ_{TH}, it is

$$\frac{\alpha_s}{\epsilon_{TH}} = \frac{\sigma A_s}{W A_p}(T_2^4 - T_1^4)$$

Dynamic Measurements [18]. With the solar simulator as the heater and the property α_s/ϵ_{TH}, one makes cooling and heating measurements. For heating, the relationship is

$$mc_p \frac{dT_s}{dt} = A_p\, \alpha_s W + P - A_s\, \epsilon_{TH} \sigma T_s^4$$

For cooling, it is

$$mc_p \frac{dT_s}{dt} = P' - A_s \epsilon_{TH} \sigma T_s^4$$

In these equations,

I = current to sample heater

R = resistance of sample heater

T_2 = equilibrium temperature of sample

T_1 = temperature of cold wall of vacuum chamber

T_s = sample temperature

A_s = total surface area of sample

A_p = projected area of sample as viewed in the direction of illumination

W = total radiant power per unit area from solar simulator

m = mass of sample

c_p = specific heat of sample per unit mass

P = incident thermal radiation power from walls of vacuum chamber

P' = incident thermal radiation for the cooldown period

t = time

Since the thermal radiation properties are dependent not only on the intrinsic material but also on the coatings, thickness, and surface condition, published numerical values for α_s and ϵ_{th} must be used cautiously. Data on the values of α_s and ϵ_{th} should include a complete statement about the coating: surface roughness, thickness, cleanliness, precise chemical composition, and a complete substrate description [22]. Current published data seldom supply sufficient descriptions of coatings, may even fail to distinguish between ϵ_{TN} and ϵ_{TH}, or may not associate a temperature with a given emittance figure. The thermal radiation properties data tabulation (Table 20-4) is meant only to be a guide in materials selection. After a coating has been selected and its method of formulation, application, and handling fixed, the α_s and ϵ_{th} properties should be measured.

In Fig. 20-20 to 20-28, the emissivities of materials and their solar absorptances are given. More details on measurement conditions, etc., are given in [12].

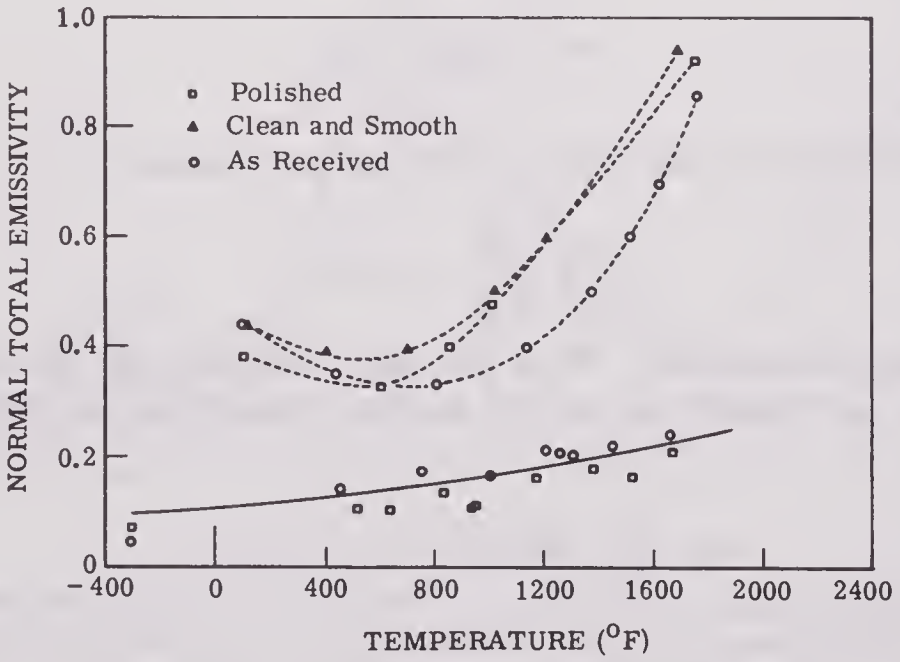


FIG. 20-20. Normal total emissivity and total solar absorptivity of Inconel X. (Note: temperatures are those to which samples had been heated before tests.)

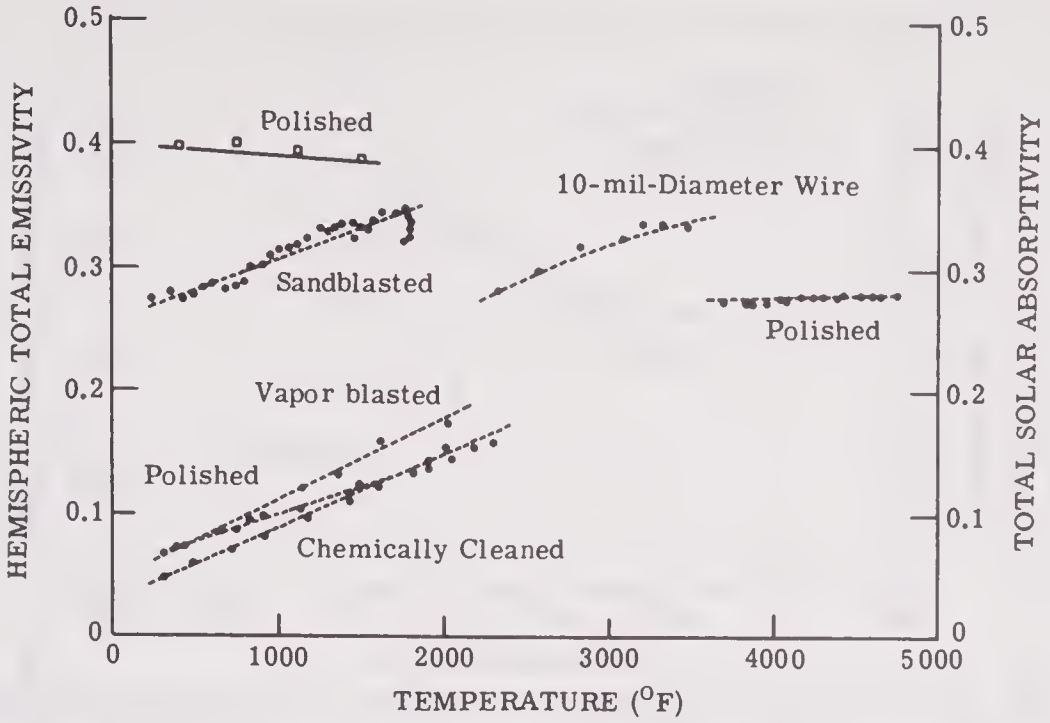


FIG. 20-21. Hemispheric total emissivity and total solar absorptivity of molybdenum.

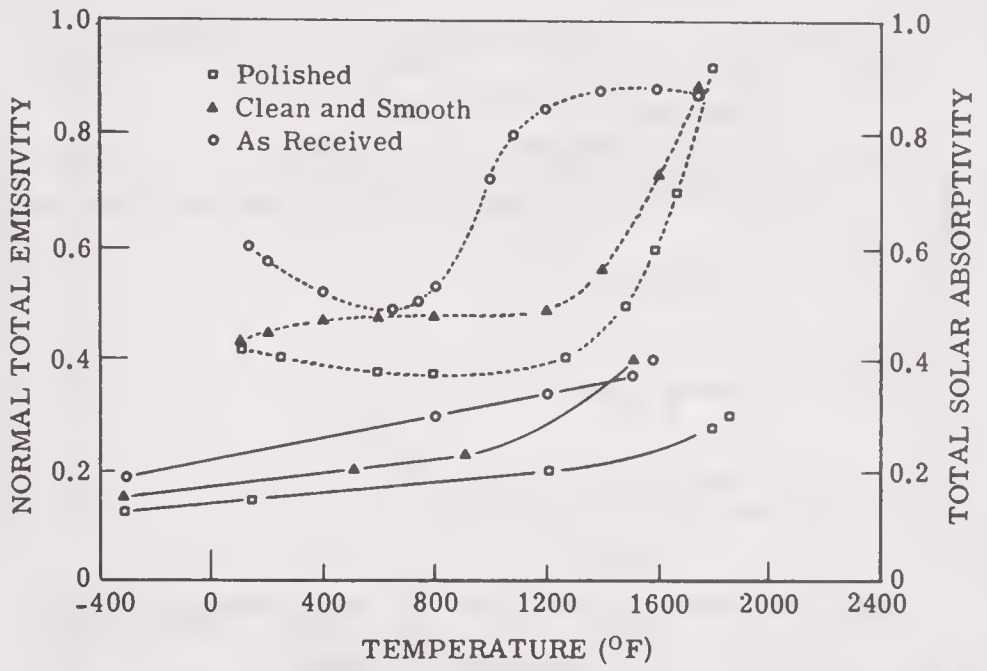


FIG. 20-22. Normal total emissivity and total solar absorptivity for K-Monel. (Note: temperatures are those to which samples had been heated before tests.)

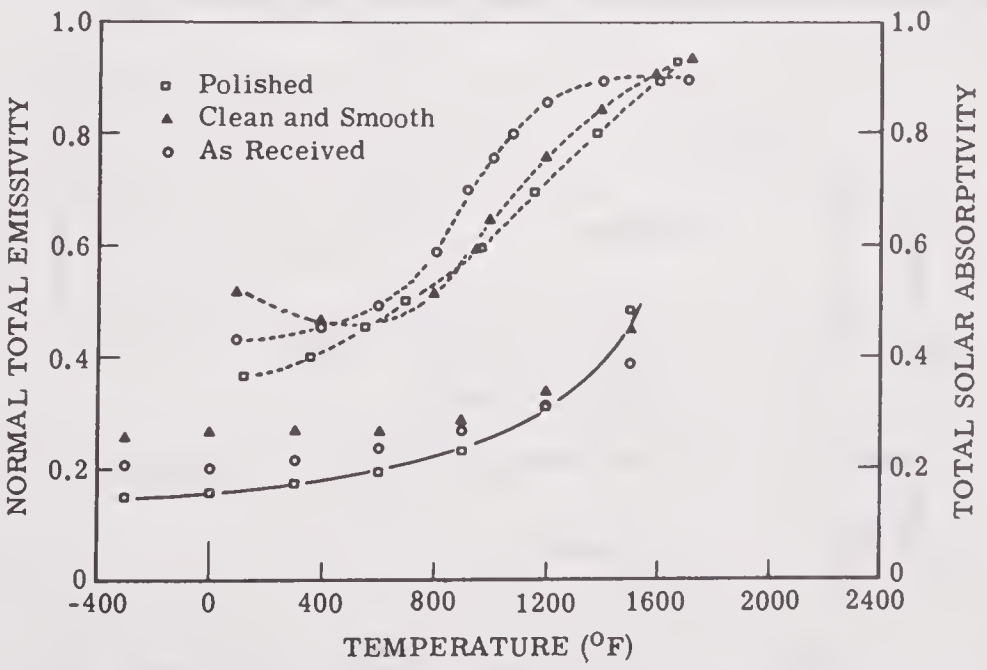


FIG. 20-23. Normal total emissivity and total solar absorptivity for 301 stainless steel. (Note: All measurements made at 100° F. Temperatures are those to which samples had been heated previous to tests.)

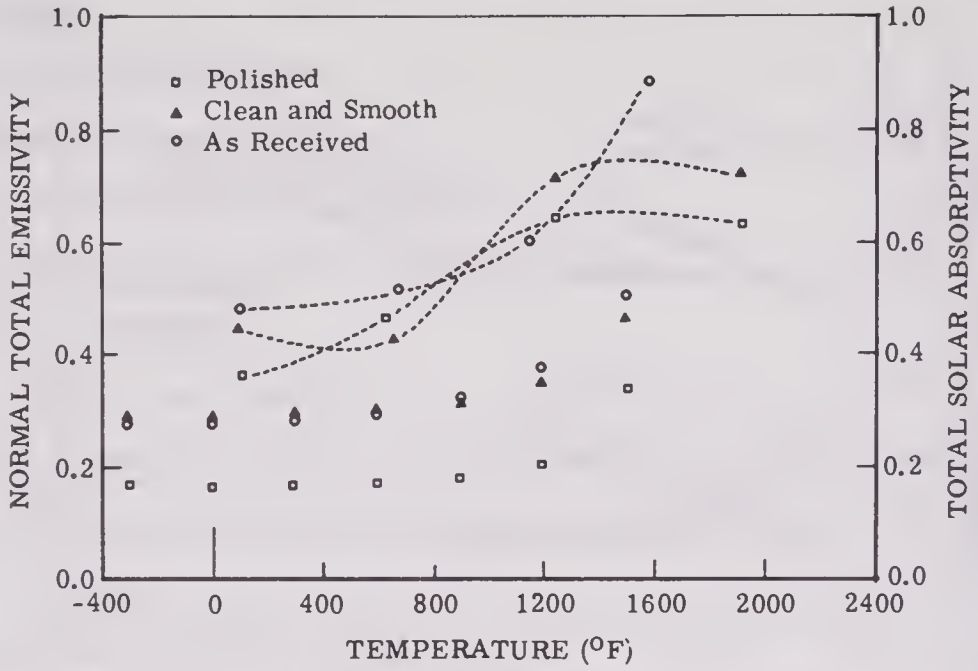


FIG. 20-24. Normal total emissivity and total solar absorptivity for 316 stainless steel. (Note: All measurements made at 100° F. Temperatures are those to which samples had been heated previous to tests.)

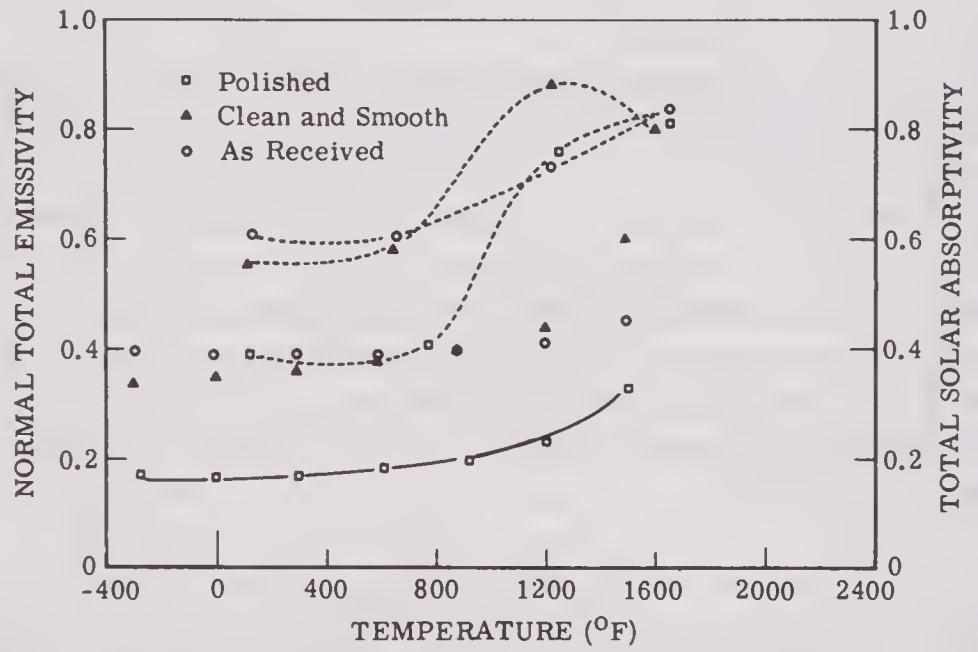


FIG. 20-25. Normal total emissivity and total solar absorptivity for 347 stainless steel. (Note: All measurements made at 100° F. Temperatures are those to which samples had been heated previous to tests.)

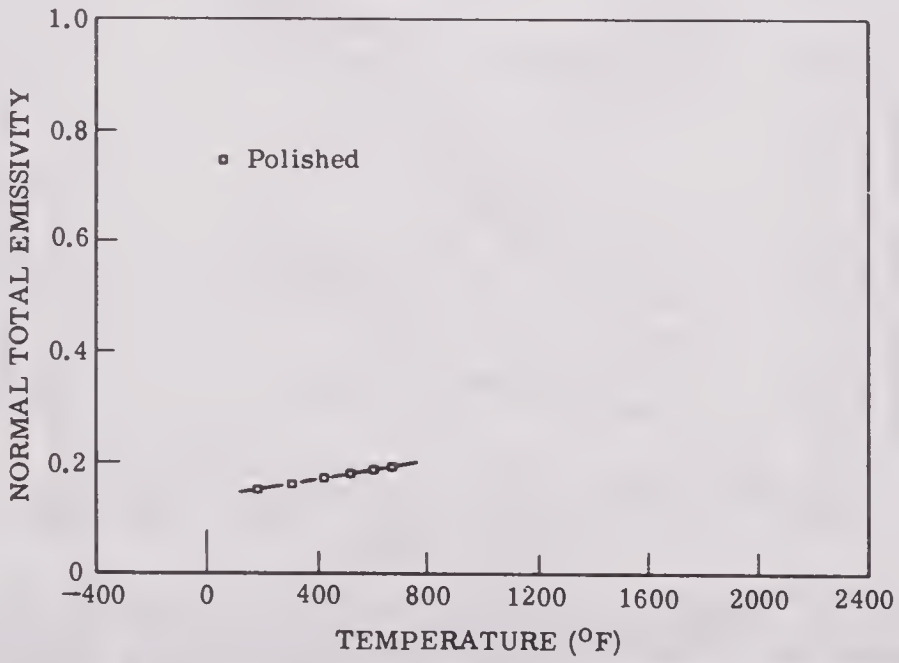


FIG. 20-26. Normal total emissivity for 18-8 stainless steel.

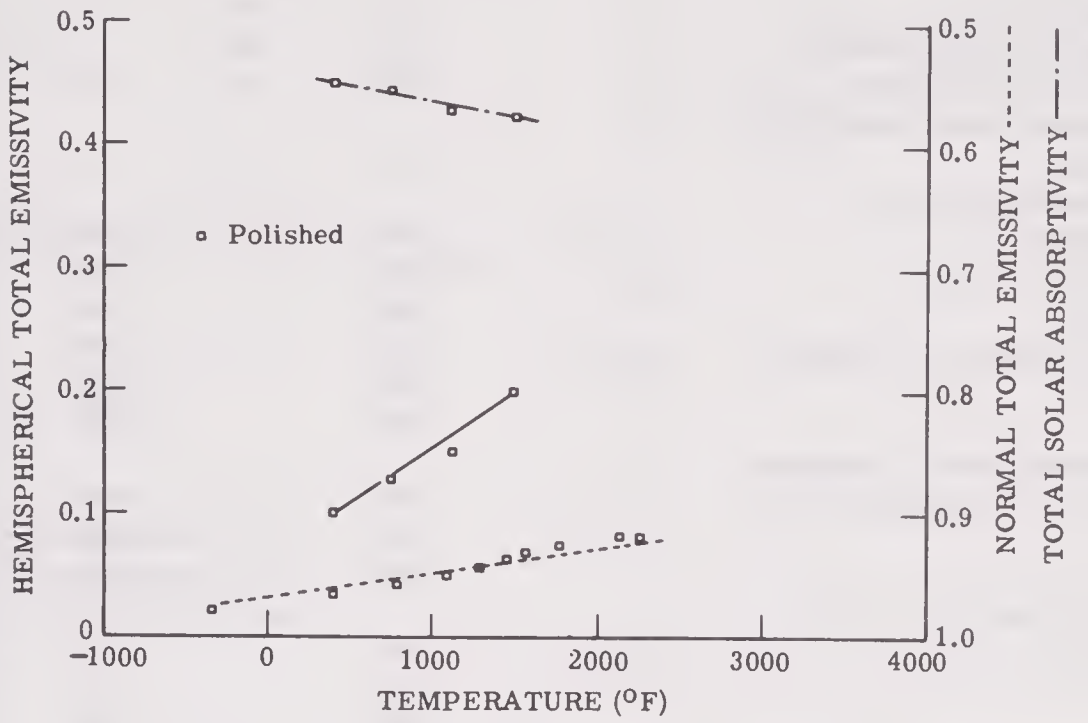


FIG. 20-27. Hemispherical total emissivity, normal total emissivity, and total solar absorptivity for tantalum.

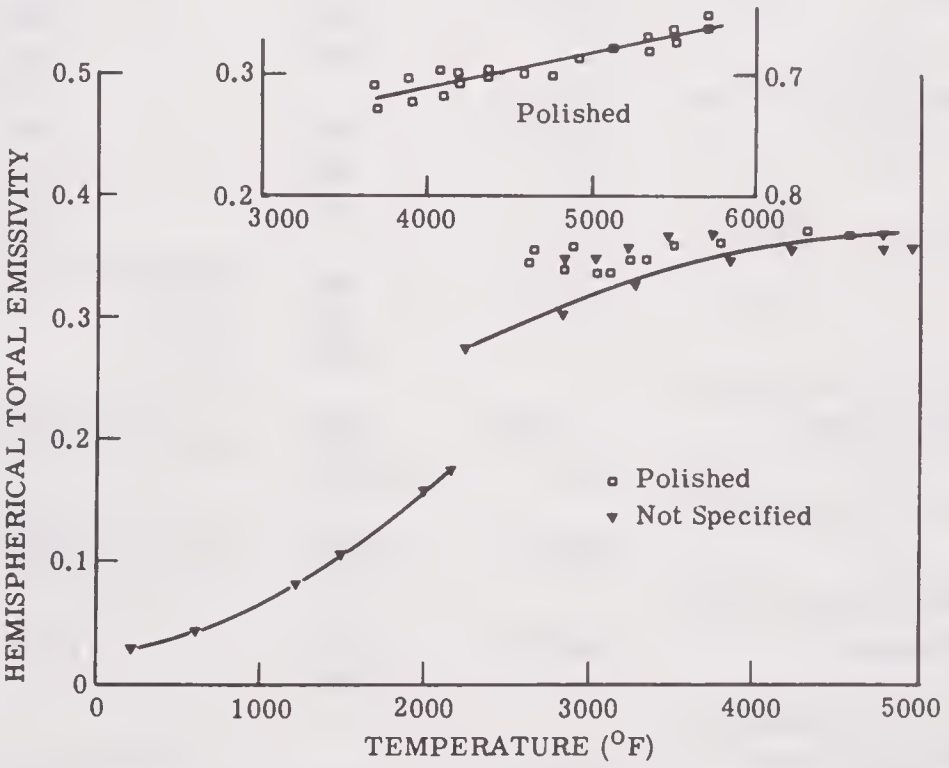


FIG. 20-28. Hemispherical total emissivity of tungsten.

TABLE 20.4. THERMAL RADIATION PROPERTIES OF MATERIALS

Material	α_s	*Temperature* (°R)	ϵ_{TN}	ϵ_{TH}	α_s/ϵ_{TH}	*Ref.*
Aluminum						
Commercial plate		671	.09			23
Commercial plate, polished		671	.05			23
Commercial plate dipped in HNO_3		671	.05			23
Commercial plate dipped in hot hydroxide		671	.04			23
2024 aluminum alloy as received	0.27	400		0.02	13.5	24
		500		0.02	13.5	
		600		0.02	13.5	
2024 aluminum alloy cleaned	0.34	400		0.06	5.666	24
		500		0.06	5.666	
		600		0.07	4.857	
2024 aluminum alloy mechanically polished and degreased	0.31	400		0.05	6.20	24
		500		0.06	6.20	
		600		0.06	5.16	
2024 aluminum alloy sand blasted	0.67	400		0.25	2.68	24
		500		0.27	2.48	
		600		0.30	2.23	
6061 aluminum alloy as received	0.41	400		0.05	8.2	24
		560		0.04	10.25	
6061 aluminum alloy chemically cleaned	0.18-0.44	400		0.03-0.11		24
		500		0.03-0.12		
		600		0.04-0.12		
6061 aluminum alloy polished and degreased	0.35	400		0.04	8.75	24
		500		0.04	8.75	
		560		0.05	7.00	
6061 aluminum alloy 120-size grit blasted	0.60	400		0.40	1.5	24
		500		0.41	1.46	
		600		0.41	1.46	
Alzak on aluminum (190 μin. thick)	0.15	410	0.79			21
		530	0.77			
		580	0.75			
Aluminum foil Reynolds wrap shiny side as received	0.19	400		0.03	6.333	24
		500		0.04	4.75	
Aluminum foil Fasson adhesive-backed	0.17	400		0.03	5.66	24
		500		0.03	5.66	
Mylar metalized with vacuum-deposited aluminum	0.20	400		0.05	4.0	24
		500		0.05	4.0	
Aluminum vacuum deposited on magnesium with standard silicone undercoat	0.13	400		0.04	3.25	24
		500		0.04	3.25	
		600		0.04	3.25	
6061 T-6 aluminum hard anodize 1 mil thick	0.923	450		0.841	1.10	21
		410	0.830			23
		460	0.842			
		560	0.863			

TABLE 20-4. THERMAL RADIATION PROPERTIES OF MATERIALS (*Continued*)

Material	α_s	*Temperature* (°R)	ϵ_{TN}	ϵ_{TH}	α_s/ϵ_{TH}	*Ref.*
Antimony						
Polished		671	0.03			23
Rolled plate		540	0.06			23
Shim stock 6 5/35		531		0.029		23
Oxidized		900	0.60			23
		671	0.60			
Beryllium						
QMV Beryllium alloy 64 RMS finish	0.70	400		0.16	4.375	24
		600		0.17	4.117	
Cadmium		540	0.02			23
Electroplate (mossy)		531		0.03		23
Chromium		540	0.08			23
Chromium plate 0.1 mil thick on 0.5 mil nickel plate on 321 stainless steel exposed to JP-4 combustion products 50 hours at 1,100°F	0.778	555		0.15	5.18	21
Cobalt		531	0.03			23
Columbium (niobium)						26
Columbium alloy (Cb-10Ti-10Mo)						26
Copper						
Black oxidized		540	0.78			23
Scraped		540	0.07			23
Commercial polish		540	0.03			23
Electrolytic, careful polish		635	0.018			23
		531		0.015		
Chromic acid dip		531		0.017		23
Polished		531		0.019		23
Liquid honed		531		0.088		23
Electrolytic polish		431	0.006			23
Mechanical polish		531	0.015			23
Pure copper, carefully prepared surface		531	0.008			23
Ebanol						
Ebanol C on copper treated 5 min at 196°F in 219°F boiling point solution	0.908	555		0.11	8.25	21
Ebanol S on steel treated 15 min in a 286°F boiling solution	0.848	555		0.10	8.49	21
Glass 3 mils thick on silicone solar cell	0.925	450		0.843	1.10	21
Gold						
0.000010-in. leaf (on glass or Lucite plastic)		531		0.063		23
0.000040-in. foil (on glass or Lucite plastic)		531		0.023		23
0.0005-in. foil (on glass or Lucite plastic)		531		0.016		23
0.0015-in. foil (on glass or Lucite plastic)		531		0.01		23
Gold plate 0.00005 in. on stainless steel (1% Ag in Au)		531		0.027		23

TABLE 20-4. THERMAL RADIATION PROPERTIES OF MATERIALS (*Continued*)

Material	α_s	*Temperature* (°R)	ϵ_{TN}	ϵ_{TH}	α_s/ϵ_{TH}	*Ref.*
Gold (*Continued*)						
Gold plate 0.0001 in. on stainless steel (1% Ag in Au)		531		0.027		23
Gold plate 0.0002 in. on stainless steel (1% Ag in Au)		531		0.025		23
Gold plate 0.0002 in. on copper (1% Ag in Au)		531		0.025		23
Gold vaporized onto 2 sides of 0.0005-in. Mylar plastic		531		0.02		23
Deep electroplated gold on aluminum		410	0.02			27
		460	0.02			
		560	0.02			
Gold plate on 7075 aluminum	0.30	400		0.03	10.0	24
		500		0.03	10.0	
Vacuum-deposited gold on buffed titanium	0.33	400		0.05	6.6	24
		500		0.05	6.6	
		600		0.05	6.6	
Vacuum-deposited gold on aluminum with resin undercoat	0.24	400		0.04	6.0	24
		500		0.04	6.0	
		600		0.04	6.0	
Graphite (crushed carbon electrodes) 16 mils thick on sodium silicate on polished aluminum	0.96	450		0.908	1.06	21
Inconel						26
Inconel foil (0.005) as received	0.55	400		0.21	2.619	24
		500		0.23	2.391	
Inconel X (Fig. 20-20)						26
Inconel X oxidized 4 hr at 1825°F in air followed by 10 hr at 1300°F in air	0.898	450		0.711	1.26	21
Iridium		531	0.04			23
Iron						26
Electrolytic		959	0.07			23
		671	0.05			
		531	0.05			
Cast iron, polished		560	0.21			23
Cast iron, oxidized		560	0.63			23
		959	0.66			23
		1460	0.76			
Iron sheet, rusted red		531	0.69			23
Iron, oxidized		2700	0.89			23
		671	0.74			
Iron, nickel alloys						26
Tinned iron sheet		535	0.064			23
Galvanized iron		657	0.07			23
Lead						
Unoxidized, polished		671	0.05			23
Gray oxidized		531	0.28			23
Oxidized at 473°K		851	0.63			23
Red lead		671	0.93			23
Lead foil 0.004 in.		531		0.036		23
Magnesium		531	0.07			23
		959	0.13			
		1460	0.18			

TABLE 20-4. THERMAL RADIATION PROPERTIES OF MATERIALS (*Continued*)

Material	α_s	*Temperature* (°R)	ϵ_{TN}	ϵ_{TH}	α_s/ϵ_{TH}	*Ref.*
Manganin, bright rolled		704	0.048			23
		531	0.076			
Molybdenum		4140	0.24			23
		3240	0.19			26
		2341	0.13			
		671	0.07			
		531	0.05			
Molybdenum as received (Fig. 20-21)	0.48	500	0.12		4.0	24
Monel						
K Monel 5700 (Fig. 20-22)						26
Nickel						26
Polished		711	0.045			23
Bright matte		711	0.041			23
Nickel foil 0.004 in.		531		0.022		23
Electrolytic		1460	0.10			23
		959	0.07			
		560	0.06			
		531	0.04			
Electroplated						
on iron and unpolished		531	0.11			23
on iron and polished		531	0.045			23
on copper		540		0.03		23
Electroless nickel (Dow Chemical Co.)	0.45	400		0.16	2.815	24
		500		0.17	2.647	
Oxidized		2700	0.85			23
		900	0.37			
Nichrome						
"Driver Harris" nichrome heater strip		460	0.11			27
		560	0.12			
		660	0.13			
Paints						
Mixing white lacquer, 4 coats, 2 coats zinc chromate primer, thickness total 0.006 in. (Ditzler)	0.21	400		0.85	0.247	24
		500		0.85	0.247	
		600		0.86	0.244	
Sherwin-Williams Kemacryl white paint No. M49WC12, 4 mils over one wash primer coat on alodined 2024-T6 aluminum		410	0.90			27
		460	0.91			
		560	0.91			
Kemacryl lacquer white No. M49WC17 4 wet coats, 1 coat pretreatment primer P40GC1 (Sherwin-Williams) 0.26		400		0.73	0.356	24
		500		0.75	0.346	
		600		0.77	0.337	
Zinc sulfide pigment, acryloid A-10 binder, pigment to volume concentration 30%, 5.5 mils over one wash primer coat on alodined 2024-T6 aluminum		410	0.90			27
		460	0.90			
		560	0.90			

TABLE 20-4. THERMAL RADIATION PROPERTIES OF MATERIALS (*Continued*)

Material	α_s	*Temperature* (°R)	ϵ_{TN}	ϵ_{TH}	α_s/ϵ_{TH}	*Ref.*
Paints (*Continued*)						
Lithofrax 72662, 10 mils thick on vapor honed titanium substrate		410	0.92			27
		460	0.93			
		560	0.93			
Lithofrax 72662, 11 mils on vapor honed 2014-T6 substrate		410	0.92			27
		460	0.92			
		560	0.93			
White paint, sodium silicate binder, zinc oxide-zinc sulfide pigmented, substrate vapor honed 2024-T3		410	0.94			27
		460	0.94			
		560	0.95			
Zinc oxide pigment, spectral grade in potassium silicate binder thickness 5 mils, substrate vapor honed 6061-T4 aluminum alloy		410	0.94			27
		460	0.95			
		560	0.95			
White epoxy resin paint "Cat-a-lac" Finch Paint and Chemical Co. No. 463-1-8 on aluminum	0.248	450		0.882	0.28	21
Skyspar A-423 color SA8818 (white) 4 wet coats over one coat epoxy primer P-323 (Andrew Paint Co.)	0.22	400		0.82	0.268	24
		500		0.83	0.265	
		600		0.84	0.261	
517-B-2 white gloss silicone (W. P. Fuller Co.) 3 wet coats	0.30	400		0.82	0.365	24
		500		0.81	0.370	
		600		0.80	0.375	
Zinc sulfide pigment in DO 808 silicone vehicle, 5-1/2 mils thick substrate vapor honed 2014-T6 aluminum alloy		410	0.90			27
		460	0.90			
		560	0.90			
Pemco No. R46H60 porcelain enamel on steel	0.22	400		0.76	0.289	24
		500		0.78	0.282	
		600		0.79	0.278	
Strong black lacquer, 4 coats, 2 coats zinc chromate primer, thickness .002 in. (Ditzler)	0.92	400		0.71	1.296	24
		500		0.74	1.243	
		600		0.75	1.227	
Kemacryl lacquer black No. M49BC12, 4 wet coats, 1 coat pre-treatment P40GC1 (Sherwin-Williams)	0.94	400		0.81	1.160	24
		500		0.83	1.133	
		600		0.84	1.119	

TABLE 20-4. THERMAL RADIATION PROPERTIES OF MATERIALS (*Continued*)

Material	α_s	*Temperature* (°R)	ϵ_{TN}	ϵ_{TH}	α_s/ϵ_{TH}	*Ref.*
Paints (*Continued*)						
Black epoxy paint 1019 2 wet coats over 1 coat 1012 primer on alodined aluminum 2014-T6 (Grumann)		410	0.90			27
		460	0.90			
		560	0.91			
Flat black epoxy resin paint "Cat-a-lac" Finch Paint and Chemical Co. No. 463-1-8 on aluminum	0.951	450		0.888	1.07	21
Cat-a-lac black epoxy paint 2 coats on 300 series stainless steel		410	0.89			27
		460	0.89			
		560	0.90			
Black epoxy paint applied on 2014-T6 alodined aluminum (Warnow)		410	0.91			27
		460	0.92			
		560	0.92			
517-B-2 flat black silicone 4 mils wet film (Fuller)	0.89	400		0.80	1.113	24
		500		0.81	1.099	
		600		0.82	1.085	
3M velvet black paint over zinc chromate primer on alodined aluminum substrate		410	0.92			27
		460	0.92			
		560	0.93			
Krylon 1602 black paint over aluminum		410	0.86			27
		460	0.86			
		560	0.88			
Aluminum leafing pigment in clear epoxy binder						
Oz. pigment/gal						
10		560	0.59			27
20		560	0.50			
40		560	0.34			
50		560	0.35			
80		560	0.45			
Palladium		531	0.03			23
Platinum		2461	0.18			23
		1460	0.10			
		959	0.06			
		671	0.05			
		531	0.03			
		522	0.016			
Quartz (fused)		531	0.932			23
Rhodium, plated on stainless steel		531	0.05			23
Rokide, flame-sprayed alumina on 410 stainless steel heated to 1300°F in air in 60 sec and held 30 sec more	0.276	450		0.801	0.34	21

TABLE 20-4. THERMAL RADIATION PROPERTIES OF MATERIALS (*Continued*)

Material	α_s	*Temperature* (°R)	ϵ_{TN}	ϵ_{TH}	α_s/ϵ_{TH}	*Ref.*
Silicon Solar Cell (International Rectifier Corp.), approximately 1 mm thick on electroless nickel plate substrate, boron-doped surface	0.938	555		0.316	2.97	21
Silver		1460	0.03			23
		671	0.025			
		531	0.022			
		492	0.02			
Solder, 50-50 on Cu		531		0.032		23
Stainless steel						
Polished		671	0.08			23
Type 301 (Fig. 20-23)						26
Type 302		760		0.048		23
Type 302 mechanically polished	0.38	400		0.17	2.24	24
		500		0.19	2.0	
		600		0.20	1.90	
Type 316 (Fig. 20-24)						21
Type 347 (Fig. 20-25)						21
Type 410 heated to 1300°F in air	0.764	555		0.130	5.88	21
Type 18-8 (Fig. 20-26)						21
Armco black oxide on type 301	0.891	450		0.746	1.19	21
Rene 41 alloy mechanically polished	0.38	400		0.17	2.24	24
		500		0.19	2.0	
		600		0.20	1.90	
Tabor						
Tabor solar collector chemical treatment of galvanized iron	0.885	555		0.122	7.25	21
Tabor solar collector chemical treatment 110-30 on nickel plated copper	0.853	555		0.049	17.4	21
Tantalum (Fig. 20-27)						26
		4140	0.26			23
		3240	0.21			
		531	0.05			
Tellurium		531	0.22			23
Titanium						
Alloy Ti-6AL-4VA						26
Alloy Ti-6AL-4VA as received	0.66	400		0.19	3.47	24
		500		0.20	3.3	
		600		0.22	3.0	
Titanium vapor coated on bright side of Reynolds wrap aluminum foil, 80 to 100 μ thick, heated 3 hr at 750°F in air	0.746	555		0.138	5.40	21
Titanium C-110M (AMS 4908) heated 300 hr at 850°F in air	0.768	555		0.198	3.88	21
Titanium C-110M (AMS 4908) heated 100 hr at 800°F in air	0.524	555		0.162	3.24	21
Titanium 75A (AMS 4901) heated 300 hr at 850°F in air	0.798	555		0.211	3.78	21
Anodized titanium	0.515	450		0.866	0.59	21
Titanox-RA 2 mils thick on black paint	0.154	450		0.885	0.17	21

TABLE 20-4. THERMAL RADIATION PROPERTIES OF MATERIALS (*Continued*)

Material	α_s	*Temperature* (°R)	ϵ_{TN}	ϵ_{TH}	α_s/ϵ_{TH}	*Ref.*
Tin						
1% indium		531	0.012			23
5% indium		531	0.017			23
Tin, 0.001 in. foil		531	0.012			23
Tungsten (Fig. 20-28)						26
Filament		4140	0.28			23
		3240	0.23			
		2340	0.15			
		1440	0.088			
		900	0.053			
		540	0.032			
Zinc		531	0.05			23

20.7. Spacecraft Thermal Balance

20.7.1. Deep Space Probes. Spacecraft traveling between planets are thermally influenced by the solar heat flux. In Fig. 20-29(*a*) the solar heat flux associated with planet distance from the sun is plotted as the ordinate [13]. The surface temperatures of an insulated flat plate normal to the sun are plotted as functions of its coating α_s/ϵ_{th} and its coordinate within the solar system.

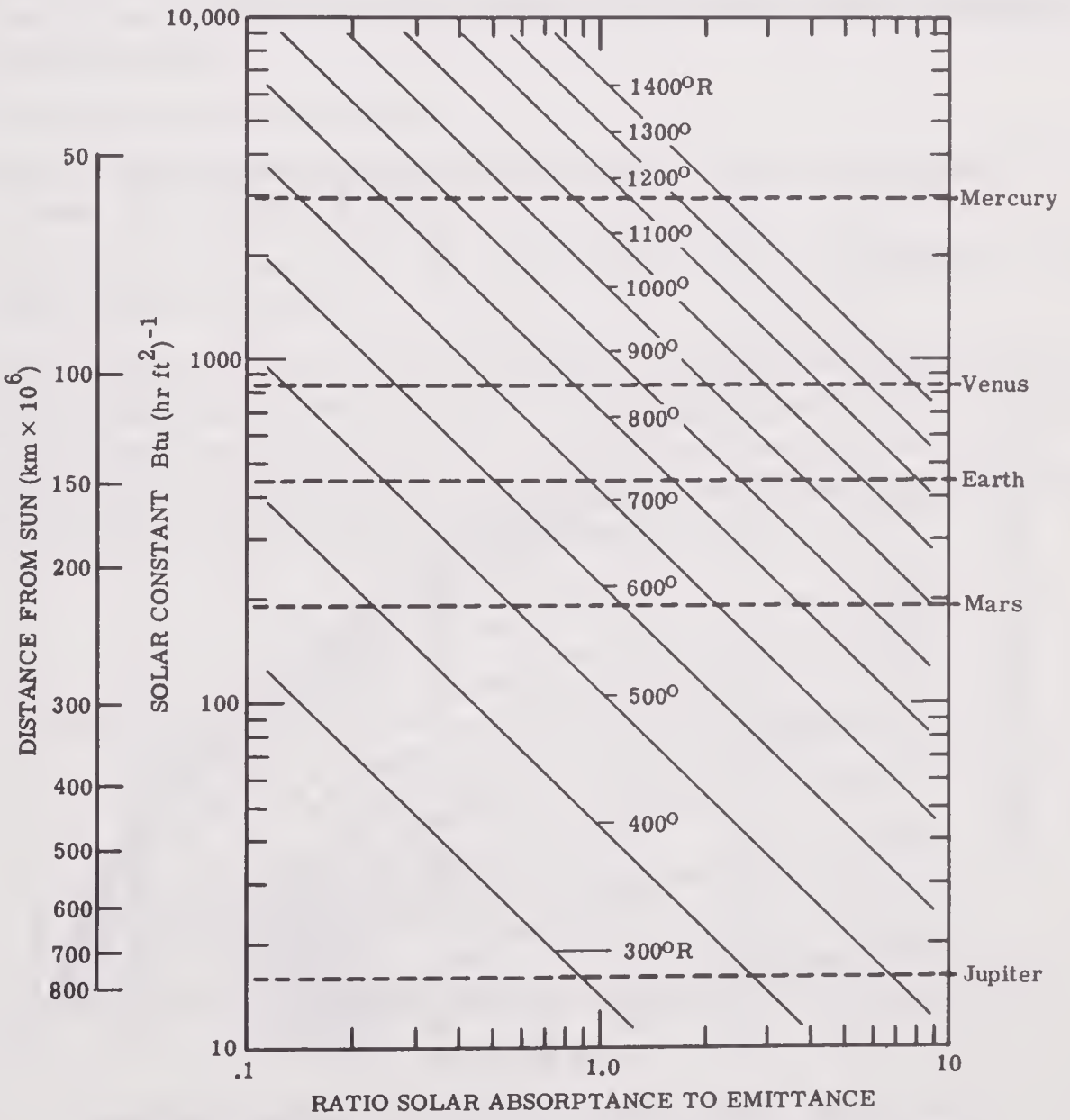


FIG. 20-29*a*. Solar heat flux as a function of distance and α/ϵ.

The heat flux (power) to any surface incremental area at an arbitrary position in space is computed by multiplying the solar power at the particular probe distance from the sun by the cosine between the surface normal and the sun line. For an average value over the spacecraft surface, this is the projected surface area to the sun.

Thus if the solar constant is written S, then the solar irradiance is:

$$H_s = S \cos \theta \tag{20-8}$$

$$H_{absorbed} = H_s \alpha_s \tag{20-9}$$

For a perfectly insulated unit area surface, the temperature, T_{sfc}, is:

$$T_{\text{sfc}} = \left(\frac{H_s \alpha_s}{\sigma \epsilon_{\text{sfc}}}\right)^{1/4}$$

The temperature of space (= 4°R) is assumed negligible for ordinary spacecraft surface temperatures.

20.7.2. Satellites. A satellite, *i.e.,* a spacecraft orbiting a planet, is irradiated from three external sources:

(1) The sun (solar radiation)

(2) Reflected sunlight off the planet (albedo radiation)

(3) The planet surface acting as a blackbody (planet radiation)

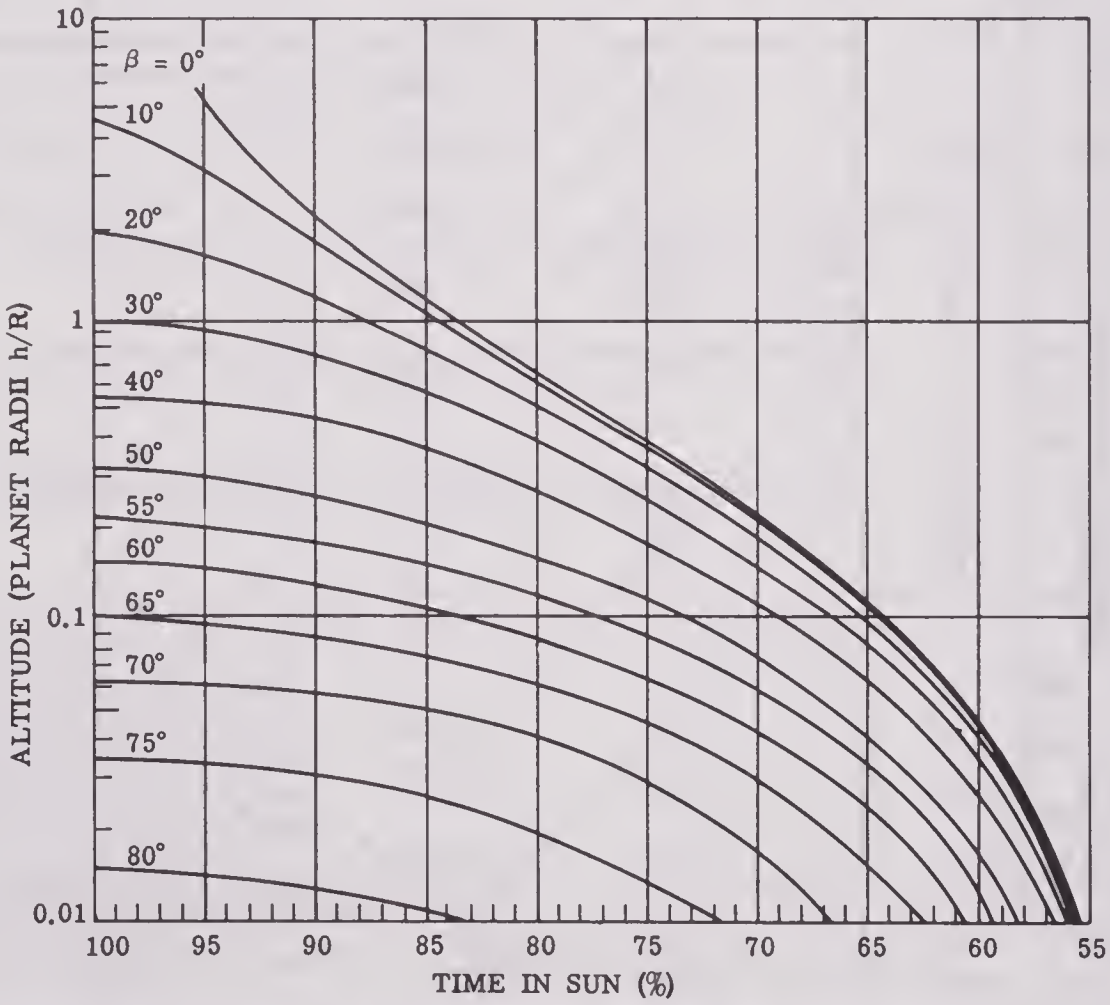


FIG. 20-29*b*. Percentage of time in sun *vs.* altitude with orbit inclination β as parameter for vehicle in circular orbit.

20.7.2.1. *Incident Solar Radiation.* The solar radiation power in a planet's orbit is obtained from Table 20-3.

In orbit, a satellite in the shadow of a planet is occulted from the sun. This shadowing, and time in the sunlight, varies according to the angle β between the satellite's orbital plane and the sun-planet plane (the ecliptic, when considering the earth; see Fig. 20-29*b*). The angle β is determined by the incident launch angle i. The orbit of an earth satellite precesses; *i.e.,* the normal to the orbital plane will generate a cone about the earth's axis at a constant half-angle of i. A 500-mi circular orbit at $i = 32°$ precesses approximately 6° per day. Because of the precession of the orbit and the rotation of the earth about the sun, the angle β varies continuously from a maximum (equatorial inclination $+i$) to a minimum (equatorial inclination $-i$). The maximum and minimum average solar heat load over the orbit are determined, respectively, when β is a maximum and when $\beta = 0°$.

Consequently, the orbital average solar heat flux incident at a point on a spatially oriented satellite's surface is

$$H_s = S \cos \theta \times (\% \text{ time in sun}) \tag{20-10}$$

where θ is the angle between the surface normal and the sun-line.

20.7.2.2. *Incident Albedo Radiation.* The parameters in the orbit geometry (Fig. 20-30) used in defining the incident albedo flux are:

θ_s = angle between the planet-sun line and the planet radius vector to the satellite

γ = angle between the satellite surface normal and the planet radius vector to the satellite surface

ν = orbit angle about the planet

ϕ_c = one of the attitude parameters, the angle of rotation about the planet radius vector to the surface normal

$\phi_c = 0$ when the normal lies in the plane containing the planet-surface vector and the planet-sun vector

ϕ_c, θ_s, and γ may vary as the satellite traverses its orbit; therefore, to obtain average albedo heat fluxes, computations of geometric factors must be made at each orbit angle ν and integrated over the orbit.

Figures 20-31 through 20-41 are used to determine the geometric factor F_a for albedo radiation incident to a sphere, cylinder, and flat plate [13].

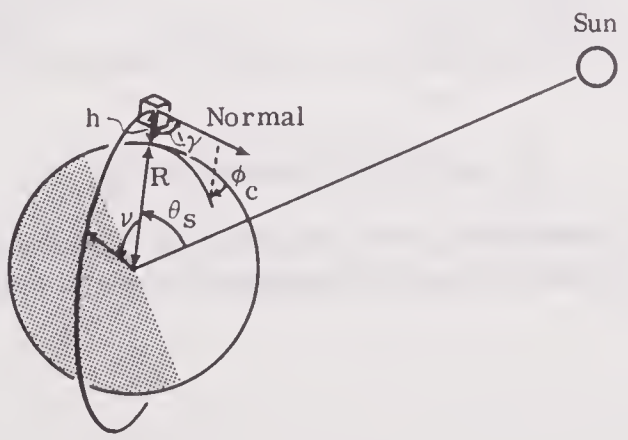


FIG. 20-30. Sun-vehicle geometry.

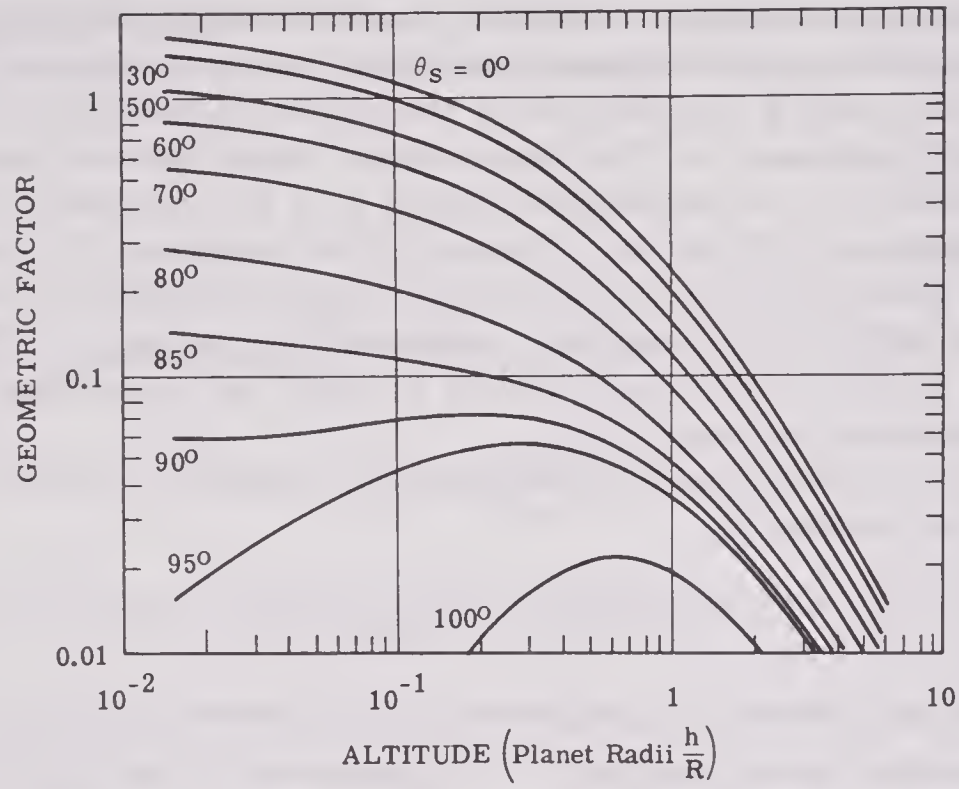


FIG. 20-31. Geometric factor for planetary albedo radiation incident to a sphere, *versus* altitude, with angle of sun as a parameter.

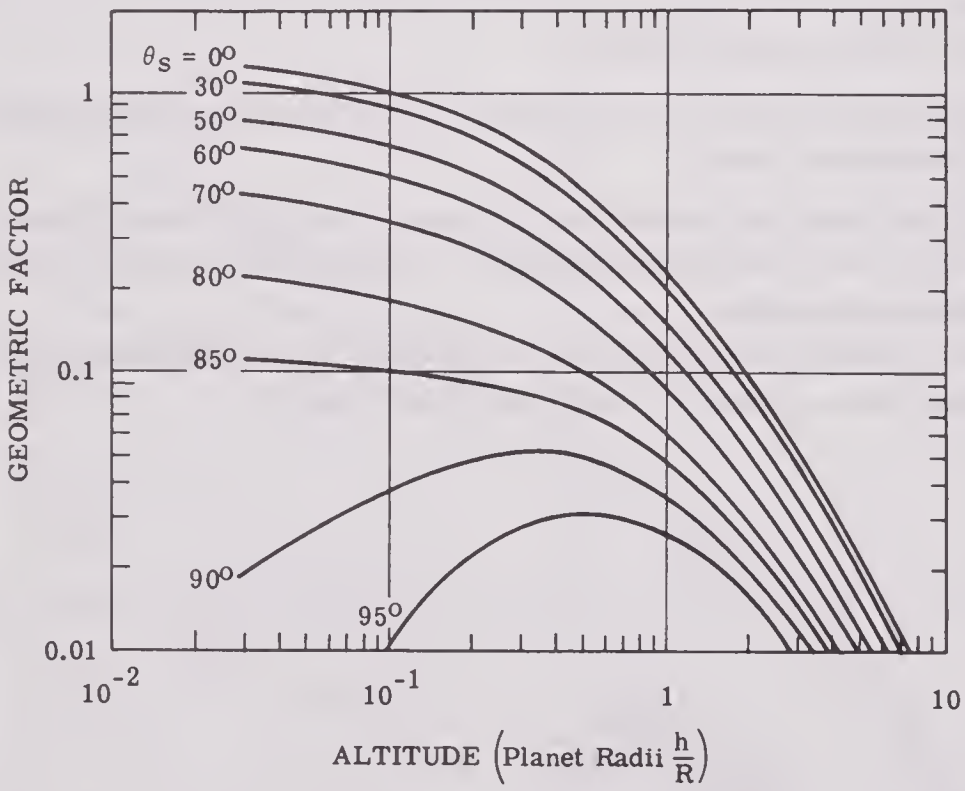


FIG. 20-32. Geometric factor for albedo to hemisphere, *versus* altitude, with angle of sun as parameter ($\gamma = 0°$, $\phi_c = 0°$).

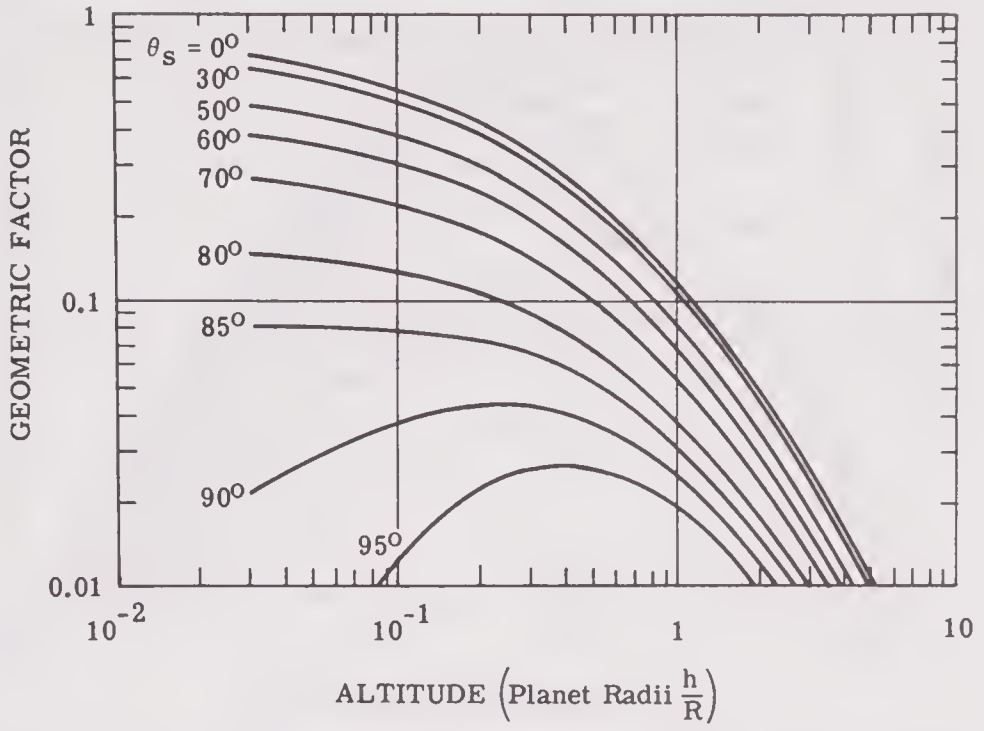


FIG. 20-33. Geometric factor for albedo to hemisphere, *versus* altitude, with angle of sun as parameter ($\gamma = 90°$, $\phi_c = 0°$).

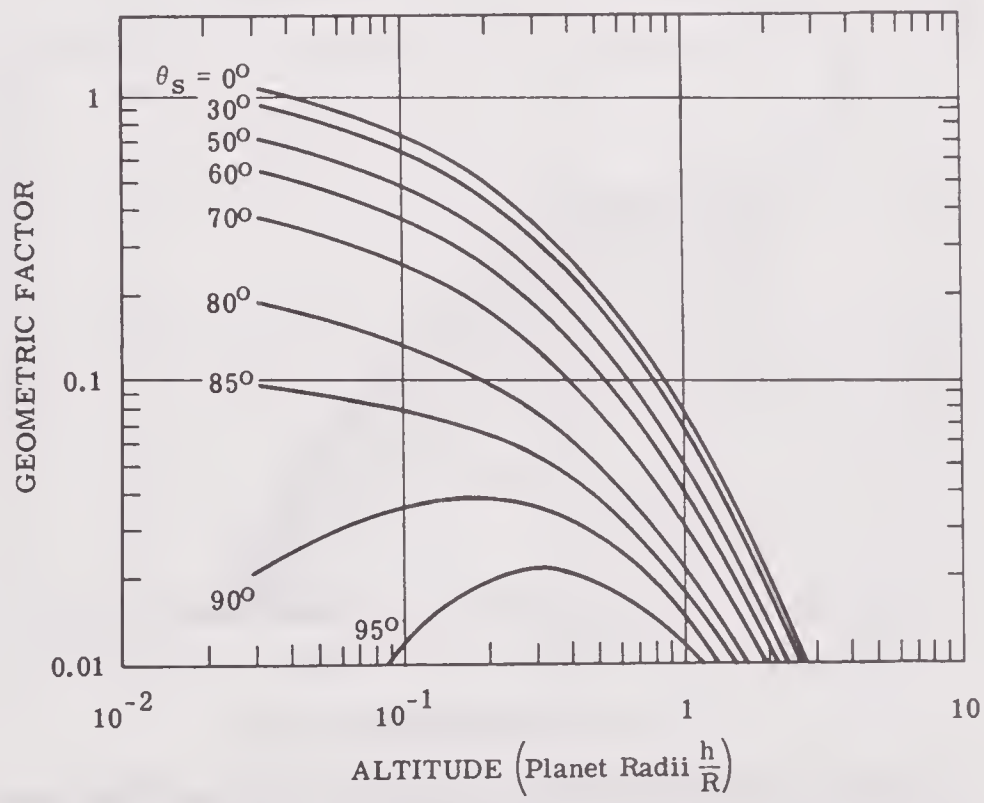


FIG. 20-34. Geometric factor albedo to cylinder, *versus* altitude, with angle of sun as parameter ($\gamma = 0°$, $\phi_c = 0°$).

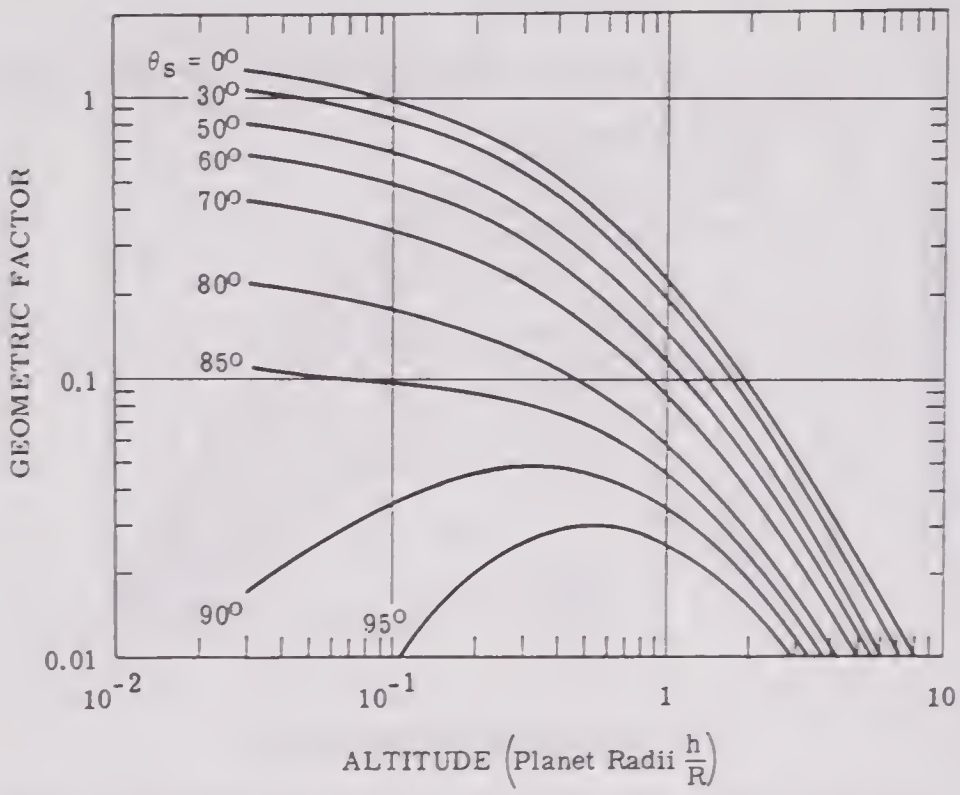


FIG. 20-35. Geometric factor for albedo to cylinder, *versus* altitude, with angle of sun as parameter ($\gamma = 90°$, $\phi_c = 0°$).

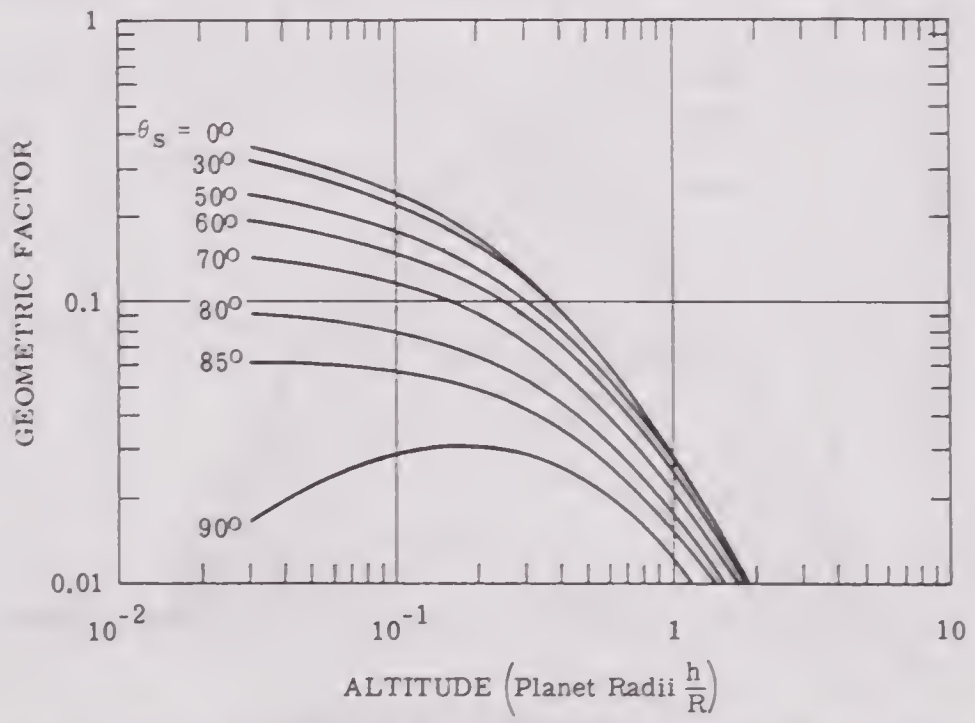


FIG. 20-36. Geometric factor for albedo to one side of flat plate, *versus* altitude, with angle of sun as parameter ($\gamma = 90°$, $\phi_c = 0°$).

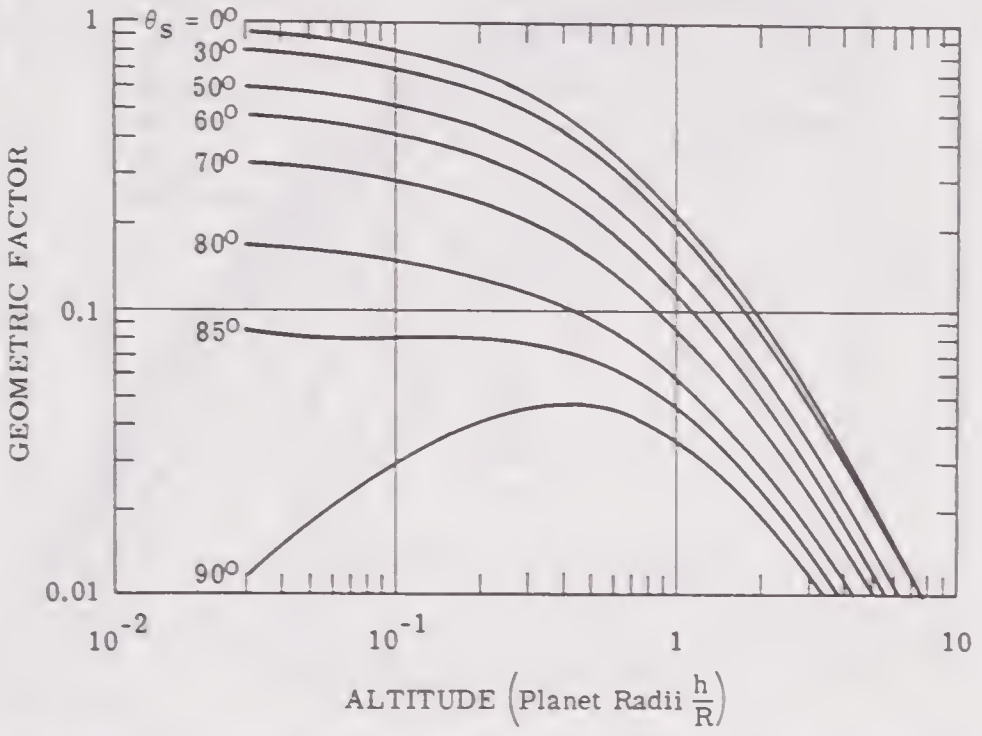


FIG. 20-37. Geometric factor for albedo to one side of flat plate, *versus* altitude, with angle of sun as parameter ($\gamma = 0°$, $\phi_c = 0°$).

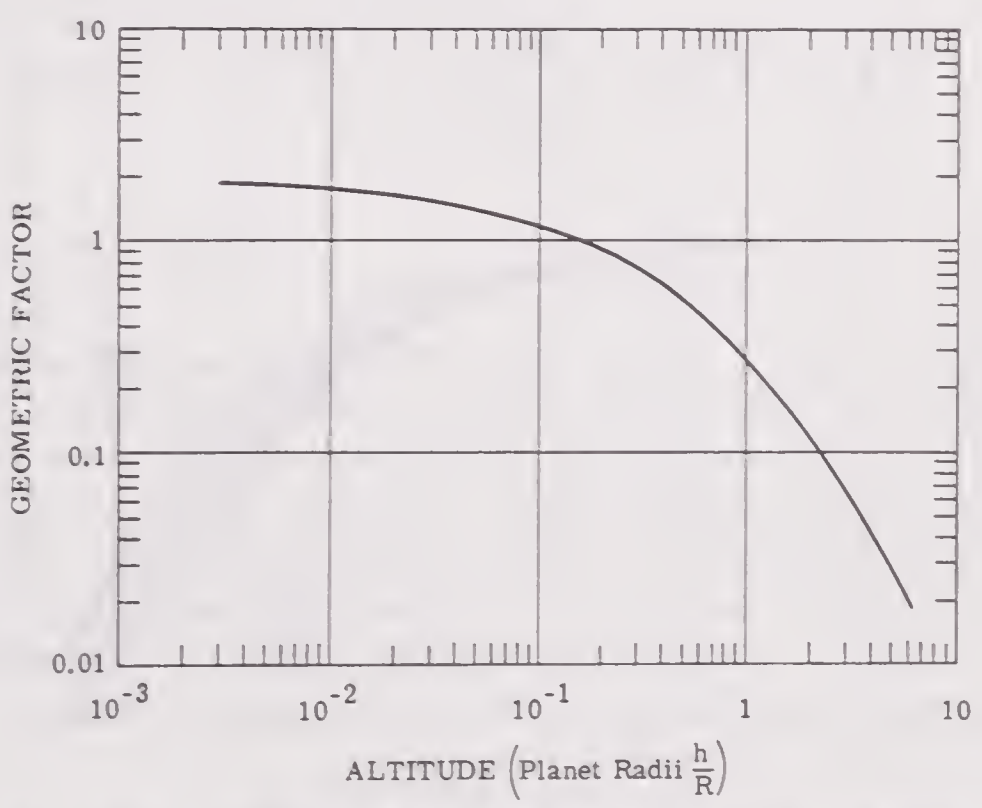


FIG. 20-38. Geometric factor for planetary thermal radiation incident to sphere, *versus* altitude.

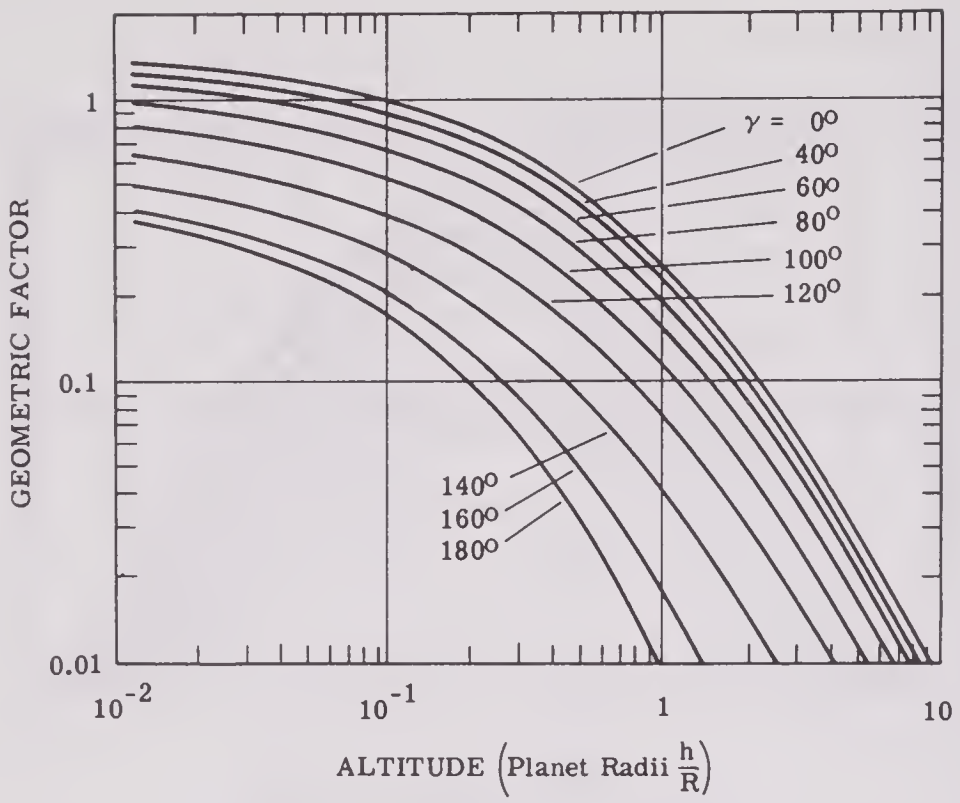


FIG. 20-39. Geometric factor for planetary thermal radiation incident to hemisphere, *versus* altitude, with attitude angle as parameter.

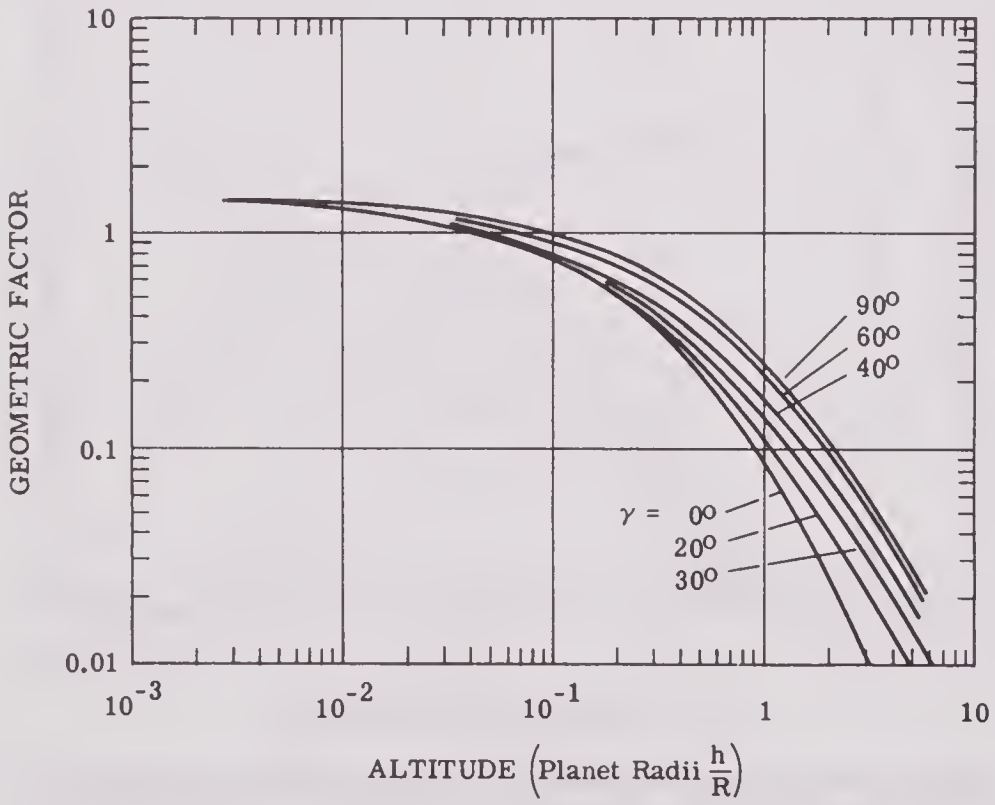


FIG. 20-40. Geometric factor for planetary thermal radiation to cylinder, *versus* altitude, with attitude angle as parameter.

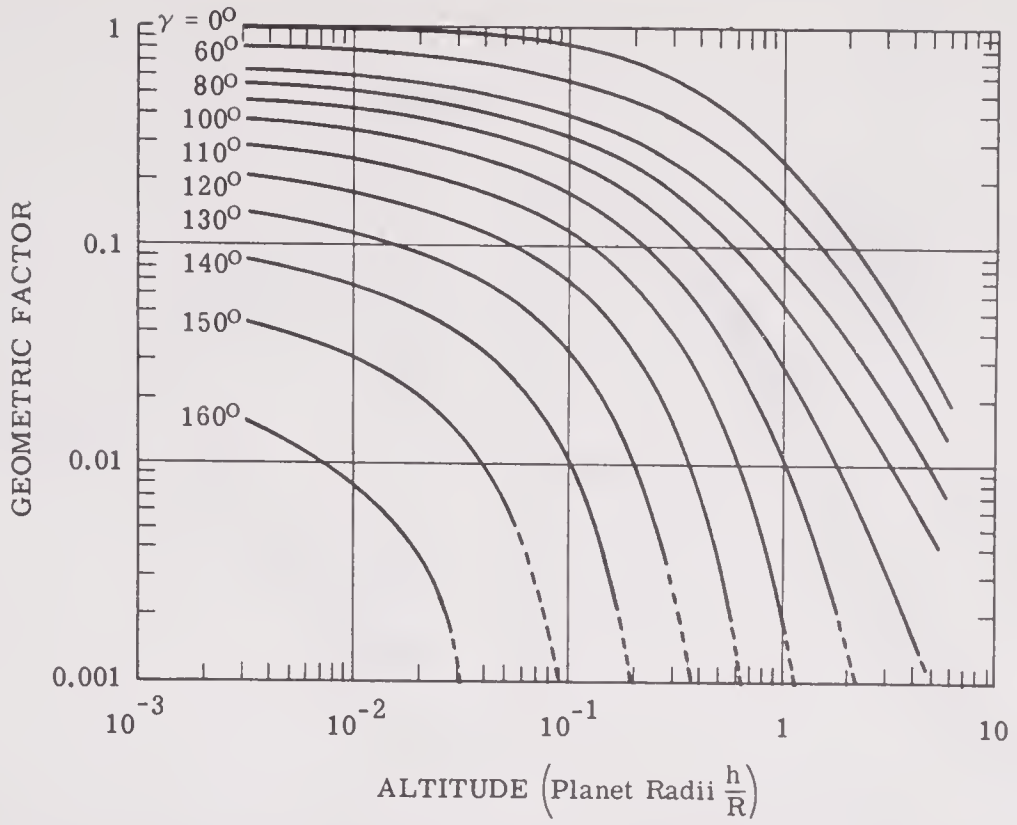


FIG. 20-41. Geometric factor for planetary thermal radiation incident to flat plate, *versus* altitude, with attitude angle as parameter.

From these, the instantaneous albedo heat flux density indicated by the prime is obtained:

$$H'_a = SaF_a$$

where a is the albedo factor (from Table 20-3).

Average heat fluxes per orbit are obtained from

$$H_a = Sa \sum_{n=1}^{n=k} F_a \Big]_{\nu_n}$$

where $\nu_n = 0°, 10°, 20°, \ldots 350°$

$n = 1, 2, 3, \ldots k$

To obtain the total power incident, $P = HA$, one can project the surface area for A in Fig. 20-42 through 20-45. For the flat-plate heat fluxes, the cosine relation is incorporated in the curves, so that the full flat-plate area can be used.

20.7.2.3. *Incident Planetary Radiation.* Similar geometric factors [13] are given below for planetary radiation to a sphere, cylinder, and flat plate using the geometry outlined above. The angle γ varies with orbit angle, and for average orbital planet flux the geometric factor must be averaged over the orbit.

$$H = H_p = \epsilon_p \sigma T_p^4 \sum_{n=1}^{n=k} \frac{F_R]_{\nu_n}}{n}$$

$$= W_p \sum_{n=1}^{n=k} \frac{F_R]_{\nu_n}}{n}$$

Generally the planet emissivity ϵ_p is assumed to be 1. T_p is obtained from Table 20-3, and $\epsilon_p \sigma T_p^4$ becomes the planet emittance W_p. The instantaneous planet heat flux is:

$$H'_p = W_p F_R$$

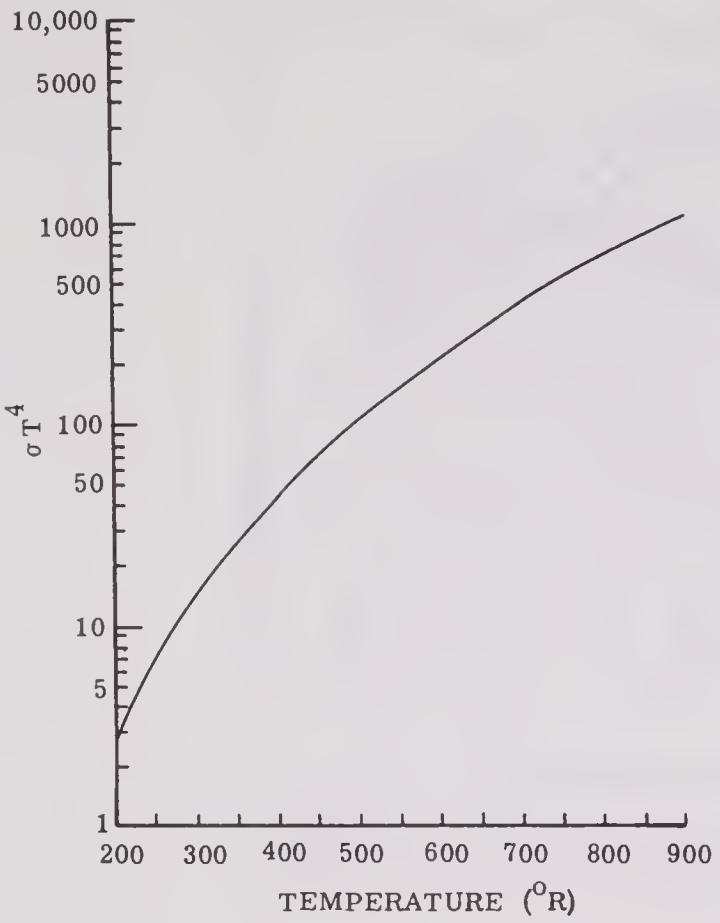


FIG. 20-42. Space radiation to an insulated thin plate, 200°-900°R.

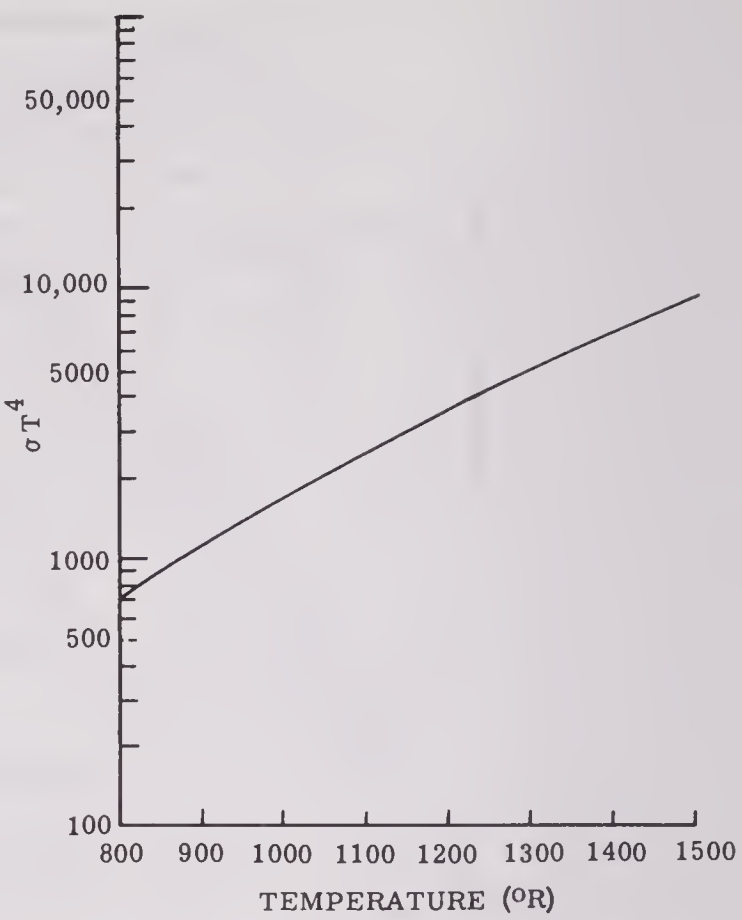


FIG. 20-43. Space radiation to an insulated thin plate, 800°-1500°R.

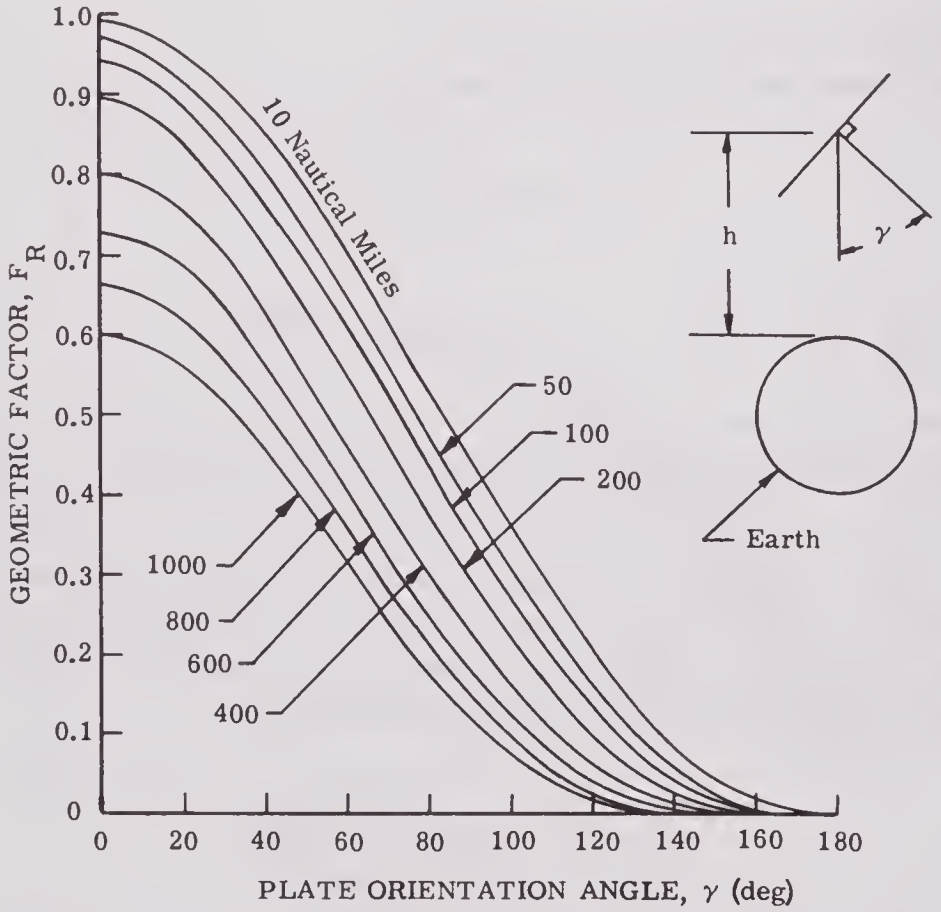


FIG. 20-44. Geometric factor for earth radiation to a flat plate, *versus* plate orientation and distance above earth surface. γ is the plate orientation angle, the true angle between the plate normal and the earth-plate line. h = plate distance above the earth's surface, in nautical miles. F_R = geometric factor for diffuse radiation from the earth to the plate. The flat plate irradiance is given by: $H_p = F_R W_p$ Btu (hr ft^2)$^{-1}$. For an average yearly albedo of 36%, H_p = 70.9 Btu (hr ft^2)$^{-1}$.

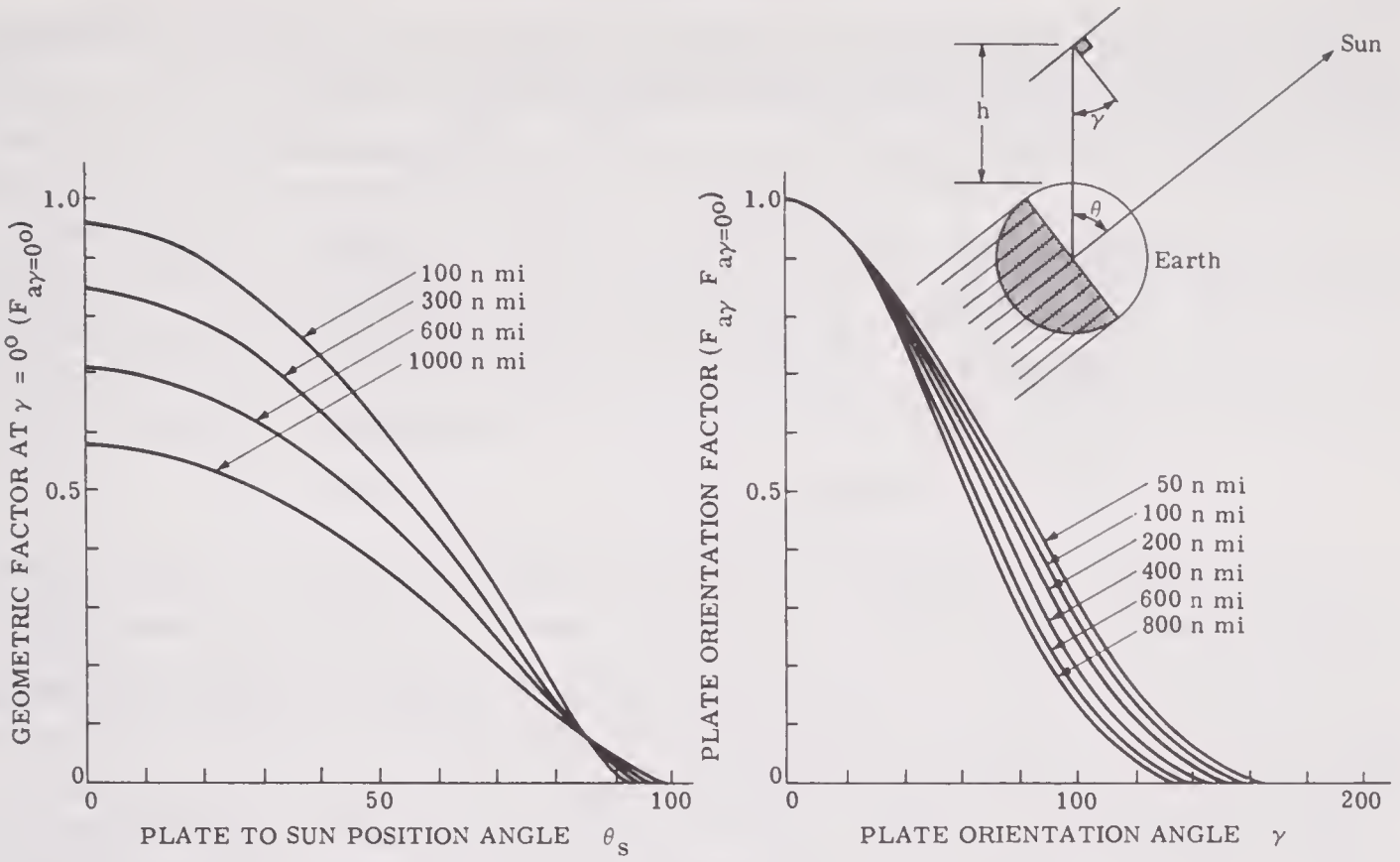


FIG. 20-45. Geometric factors for earth albedo to a flat plate, *versus* plate orientation and distance above earth's surface. $F_{a\gamma=0}$ is the geometric factor from the earth to a plate parallel to the earth surface. $F_{a\gamma}/F_{a\gamma=0}$ is the plate orientation correction factor as a function of altitude. γ is the true angle between plate normal and earth-plate line. θ_s is the true angle between earth-plate and earth-sun lines. h is the plate distance above earth's surface, in nautical miles. The flat plate irradiance is given by: $H = F_{\gamma=0}\ (F_\gamma/F_{\gamma=0})H_a$. For the average yearly earth albedo of 36%, $H_a = 159.3$ Btu (hr ft^2)$^{-1}$.

20.7.2.4. *Total Incident Space Radiation.* If one sums the heat fluxes of solar, albedo, and planet radiation, the instantaneous heat flux to a satellite's surface at a point is calculated from:

$$H' = S \cos \theta + aSF_a + W_pF_R$$

as a function of ϕ_c, θ, θ_s, and γ.

The orbital average heat flux to a surface is calculated from:

$$H = S \cos \theta + aS \sum_{n=1}^{n=k} \frac{F_a]_{\nu_n}}{n} + W_p \sum_{n=1}^{n=k} \frac{F_R]_{\nu_n}}{n} = H_s + H_a + H_p$$

Once the irradiances from space are known, along with the corresponding spacecraft surface thermal properties of α_s, α_a, and ϵ_{sfc}, the surface temperature (T_{sfc}) of a perfectly insulated material may be computed from:

$$\epsilon_{sfc}\sigma T_{sfc}^4 = H_s\alpha_s + H_a\alpha_a + H_p\alpha_p$$

Generally, the absorption of the materials for albedo radiation is the same as the absorption for the solar spectrum, and for the case where the planet temperature and skin temperature are similar, the surface absorption for planet radiation (α_p) is the same as the surface emissivity; *i.e.*, $\alpha_p = \epsilon_{sfc}$. The above equation then becomes

$$\sigma T_{sfc}^4 = (H_s + H_a)\frac{\alpha_s}{\epsilon_{sfc}} + H_p$$

20.7.3. Thermal Design. In general, the thermal coating is the best possible for the maintenance of life, electronic equipment, and structural integrity within specified temperature requirements.

Refer to Chapters 8 and 12 for the conditions under which materials and detectors operate. Other components have the following temperature ranges:

Component	*Temperature* (°F)
Batteries	40-100
Computers	25-100
Electronics	0-130
Gyros	60-160
Control gas storage	0-130

To satisfy these requirements, the spacecraft thermal coupling to space must be designed to carry the internal power load and yet satisfy the limitations of temperature.

20.7.3.1. *Passive Control.* These conditions can be satisfied by a passive system if the orientations and equipment power dissipations are known so that an α_s/ϵ_{th} can be chosen for operation within the temperature requirements.

Figure 20-46 has been drawn to show the variation in equipment temperature as a function of its power dissipation and of thermal coating α_s and ϵ_{th}. The figure incorporates solar, earth albedo, and earth radiation to a flat plate for a 500-mi circular

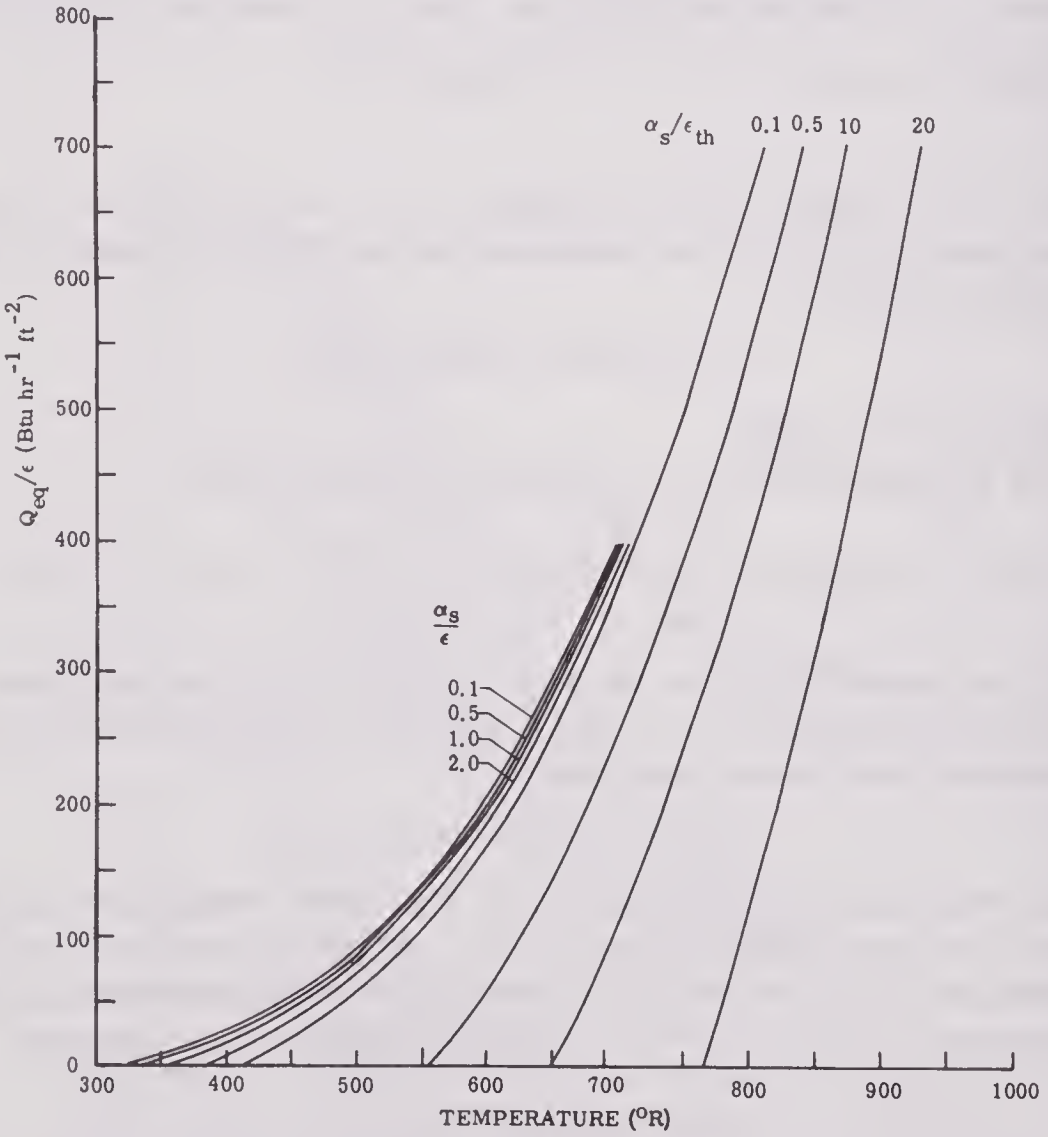


FIG. 20-46. Temperature of equipment heat sink *versus* equipment power dissipation P as a function of heat-sink thermal coating α_s/ϵ_{th} for a flat plate. The four right-hand curves are for heat sink normal to sun, 65% time in sun, $H_s = 286$, $H_a = 1.3$, $H_p = 21.3$ Btu (hr ft²)$^{-1}$. The left-hand curves are for equipment heat sink parallel to the plane of the ecliptic with the plate facing north, $H_s = 0$, $H_a = 10.53$, $H_p = 15.39$ Btu (hr ft²)$^{-1}$. 500-mile earth orbit (circular) in plane of ecliptic.

earth orbit. Five sides of the equipment box are assumed to be insulated; the front surface (heat sink), which faces the space environment, is not. The maximum environmental heat input assumes a heat sink normal to the sun, and the minimum heat input assumes the surface parallel to the sun's rays.

If one knows the power dissipation per square foot of heat-sink area, the thermal-coating properties of α_s and ϵ_{th} can be selected, so that temperature requirements are not exceeded, with the unit in or out of the sunshine.

The heat balance is calculated from the following equation, where the subscript "eq" means "an equipment parameter":

$$(H_s + H_a)\frac{\alpha_s}{\epsilon_{th}} + H_p + \frac{H_{eq}}{\epsilon_{th}} = \sigma T^4$$

20.7.3.2. *Active Control.* When conditions cause the temperature to fall outside the required range, an active system may be employed.

Mechanical means such as pinwheels or louvres may be used to vary the α_s/ϵ_{th} of the heat sink. Other methods such as electrical-thermostat heat controls or variable conductance systems alter the internal heat flow to the heat-sink surface and in like manner create a stabilizing effect on component temperatures.

20.8. Testing

Testing of spacecraft can involve either cold-wall vacuum operation, cold-wall vacuum plus solar simulation, or cold-wall vacuum plus solar simulation and simulation of albedo and earth radiation.

A typical test facility designed to simulate true space conditions of solar, albedo, and earth radiation, plus vacuum and cold space is in use by the National Aeronautical Space Administration at Langley Field, Virginia [28].

References

1. A. G. Worthing and D. Halliday, *Heat,* Wiley, New York, p. 435 (1948).
2. W. N. Harrison and J. C. Richmond, *Standardization of Thermal Emittance Measurements,* WADC 59-510, Aeronautical Systems Division, Dayton, Ohio (1960).
3. J. Ashbrook, G. F. Schilling, and T. E. Sterne, *Glossary of Astronomical Terms for the Description of Satellite Orbits.*
4. E. Fried, *Thermal Joint Conductance in a Vacuum,* ASME, 63-AHGT-18, Aviation and Space, Hydraulic, and Gas Turbine Conference and Products Show, Los Angeles (March 3-7, 1963).
5. E. Fried and F. Costello, *ARS Journal* **32,** 237 (1962).
6. D. O. Kern, *Process Heat Transfer,* McGraw-Hill, New York, pp. 80-82 (1950).
7. W. R. Morgan and D. C. Hamilton, *Radiant-Interchange Configuration Factors,* NACA TN 2836, Purdue University, LaFayette, Ind. (1952).
8. M. Jakob, *Heat Transfer,* Vol. I, Wiley, New York (1949).
9. W. H. McAdams, *Heat Transmission,* 3d ed, McGraw-Hill, New York (1954).
10. F. S. Johnson, *J. Meteorol.,* **11,** pp. 431-439 (1954).
11. L. D. Jaffe and J. B. Rittenhouse, *ARS Journal,* **32,** p. 336 (March 1962).
12. K. A. Ehricke, *Space Flight; Principles of Guided Missile Design,* Van Nostrand, Princeton, N.J., pp. 118-119 (1960).
13. J. C. Ballinger, J. C. Elizalde, and E. H. Christensen, "Thermal Environment of Interplanetary Space," 344B, SAE National Aeronautic Meeting, New York (1961).
14. C. G. Claus and J. B. Singletary (eds.), *Space Materials Handbook,* Lockheed Missiles and Space Co., Sunnyvale, Calif., AF 04(647)-673, pp. 32, 47, 48, 128, 129, 183 (1962).
15. F. Sartwell, *National Geographic,* **123,** p. 733 (1963).
16. W. Nordberg *et al., Preliminary Results of Radiation Measurements from the TIROS III Meteorological Satellite,* TND-1338, National Aeronautics and Space Administration, Washington, D.C. (1962).
17. S. B. Nicholson and E. Pettit, "Lunar Radiation and Temperature," p. 392, Carnegie Institution, Washington, D.C.

18. W. B. Fussell, J. J. Triolo, and J. H. Henninger, *A Dynamic Thermal Vacuum Technique for Measuring the Solar Absorptance and Thermal Emittance of Spacecraft Coatings*, TND-1716, National Aeronautics and Space Administration, Washington, D.C. (1963).
19. D. K. Edwards *et al.*, *J. Opt. Soc. Am.*, **51**, pp. 1279-1288 (1961).
20. R. V. Dunkle *et al.*, *Heated Cavity Reflectometer for Angular Reflectance Measurements*, Academic Press, New York, pp. 541-562 (1962).
21. D. K. Edwards *et al.*, *Basic Studies on the Use and Control of Solar Energy*, Annual Report NSF-G-9505, Report 60-93, University of California, Los Angeles (Oct. 1960).
22. R. E. Gaumer, "Determination of the Effects of Satellite Environment on the Thermal Radiation Characteristics of Surfaces," Paper 339C, SAE National Aeronautic Meeting, New York.
23. *American Institute of Physics Handbook*, McGraw-Hill, New York, pp. 68-71 (1957).
24. R. E. Gaumer and L. A. McKellar, *Thermal Radiative Control Surfaces for Spacecraft*, LMSD-704014, Lockheed Missiles and Space Division, Sunnyvale, Calif. (March 1961).
25. J. H. Weaver, *Anodized Aluminum Coatings for Temperature Control of Space Vehicles*, Technical Memorandum ASCN 62-9, Aeronautical Systems Division, Dayton, Ohio (March 1962).
26. W. D. Wood and C. F. Lucks, *Thermal Radiative Properties of Selected Materials*, DMIC Report 177, Vol. I, Defense Metals Information Center, Battelle Memorial Institute, Columbus, Ohio (Nov. 15, 1962).
27. Grumman Aircraft Engineering Corporation data.
28. L. G. Clark and K. A. Laband, "Orbital Station Temperature Control," *Astronautics*, **7**, pp. 40-41, (Sept. 1962).

Bibliography

Anderson, D. L., and G. J. Nothwang, *Effects of Sputtering with Hydrogen Ions on Total Hemispherical Emittance of Several Metallic Surfaces,* NASA TND-1646, National Aeronautics and Space Administration, Washington (January 1963).

Clauss, F. J., (ed.), *First Symposium on Surface Effects on Spacecraft Materials*, John Wiley & Sons, New York, 1960.

F. G. Cunningham, *Power Input to a Small Flat Plate from a Diffusely Radiating Sphere with Application to Earth Satellites: The Spinning Plate,* NASA TND-1545, National Aeronautics and Space Administration, Washington (February 1963).

Hastings, E. C., Jr., R. E. Turner, and K. C. Speegle, *Thermal Design of Explorer XIII Micrometeoroid Satellite*, NASA TND-1001, National Aeronautics and Space Administration, Washington (May 1962).

Hayes, R. J., *Quarterly Progress Report Determination of the Emissivity of Materials*, PNA-2163, Pratt and Whitney Aircraft Division of United Aircraft Corp., Hartford, Conn.

Hembach, R. J., L. H. Hemmerdinger, and A. J. Katz, *Heated Cavity Reflectometer Modifications*, ADR. 04-03-62.2, Grumman Aircraft Engineering Corp., Bethpage, L.I., N.Y. (July 1962).

Hemmerdinger, L. H., *General Description and Operation of Grumman Spectrophotometers*, PM-27, Grumman Aircraft Engineering Corp., Bethpage, L.I., N.Y. (June 1961).

Hemmerdinger, L. H., "Thermal Design of the OAO," *J. Spacecraft Rockets*, **1,5,** (Sept.-Oct. 1964).

Henderson, R. E., and D. L. Dresser, "Solar Thermionic Space Power Systems," Paper 350C, SAE National Aeronautic Meeting, New York (1961).

Katz, A. J., *Determination of Thermal Radiation Incident upon The Surfaces of an Earth Satellite in an Elliptical Orbit*, XP12-20, Grumman Aircraft Engineering Corp., Bethpage, L.I., N.Y. (May 1960).

Sadler, R., and L. H. Hemmerdinger, *The Emissometer—A Device for Measuring Total Hemispherical Emittance,* PM-001-33, Grumman Aircraft Engineering Corp., Bethpage, L.I., N.Y. (Jan. 1963).

Samela, D. A., "Thermal Analysis of the Echo II Canister Assembly," ASME, 63-AHGT-54, Aviation and Space, Hydraulic, and Gas Turbine Conference and Products Show, Los Angeles, Calif. (March 3-7, 1963).

Weaver, J. H., *Anodized Aluminum Coatings for Temperature Control of Space Vehicles*, ASD-TDR-62-918, Aeronautical Systems Division, Dayton, Ohio (Feb. 1963).

Radiation Heat Transfer Analysis of Space Vehicles, ASD-TR-61-119, Part II, Aeronautical Systems Division, Dayton, Ohio (Sept. 1962).

Chapter 21
AERODYNAMIC INFLUENCES ON INFRARED SYSTEM DESIGN

Lawrence D. Lorah and Eugene Rubin

Mithras, Incorporated

CONTENTS

21.1. Introduction 826
21.2. Infrared Windows 826
21.2.1. Window Requirements 826
21.2.2. Window Heating 827
21.2.3. Typical Results. 832
21.2.4. Window Transmission and Radiation at Elevated Temperatures . 833
21.2.5. Method of Alleviating Hot-Window Problems. 836
21.3. Refraction by the Field of Flow 839
21.3.1. Index of Refraction of Air 839
21.3.2. Shock-Wave Effects. 839
21.3.3. Boundary-Layer Effects 846
21.3.4. Fluctuating-Boundary-Layer Effects 848
21.4. Radiation from Heated Air 851

21. Aerodynamic Influences on Infrared System Design

21.1. Introduction

This chapter discusses the following major problems encountered by supersonic missiles and aircraft: first, the influence of high-speed flight on the window covering the infrared equipment; second, the optical influence exerted by the aerodynamic flow field around the vehicle; and third, the background radiation produced by the heated air surrounding the vehicle.

The discussions in Sec. 21.2 and 21.3 are restricted to supersonic flight over the Mach number range of about 1.5 to 7.0. At speeds below this, the problems considered here do not exist and system design requires only minor structural differences from those in surface-based equipment. At speeds above Mach 7 in atmospheric flight the assumption of perfect gas flow begins to break down and the chemistry of the air must be considered. The nature of the aerodynamic effects on infrared systems is the same at these higher speeds, but the details of predicting these effects are much more complex and beyond the scope of this discussion. In this chapter the ratio of specific heats for air is taken as 1.4 and the numerical coefficients are evaluated accordingly. In Sec. 21.4 the chemistry of the air is a factor and the limitation to Mach numbers below 7 no longer exists.

21.2. Infrared Windows

Infrared windows are basically protective coverings over the cavities or bays that house infrared or optical equipment. Generally, they are not an integral part of the functional scheme of the equipment and any change they make in radiation falling on the system produces a degradation of system performance. Satisfactory windows for ground-based equipment or low-speed aerial equipment can generally be designed by careful selection of a window material from those of Chapter 8 and then applying common optical and structural design techniques.

21.2.1. Window Requirements. The primary purpose of the window, whether it be a missile nose dome or a flat reconnaissance window well back along the aircraft fuselage, is to protect infrared equipment from the aerodynamic loads and severe heating rates encountered in supersonic flight. While performing this basic function, the window must meet the following additional requirements:

1. The window must introduce negligible distortion in the image formed by the optical system. This consideration nearly always results in a window which is a portion of a sphere or is made with one or more flat surfaces. Other shapes (cones, ogives, etc.) give large refraction errors and resolution losses unless the window is made very thin (see Chapter 8). However, windows that are too thin cannot carry the necessary structural loads.
2. The window must transmit radiation in the appropriate wavelengths without gross attenuation. The usual window material selection is complicated in this case by the fact that many materials lose transparency as their temperature increases, particularly near the long-wavelength cutoff (see Chapter 8).

3. The window must radiate a negligible amount of energy in the wavelength region of system operation. Even though the window emissivity is low in these wavelengths (by virtue of requirement 2) the window proximity to the sensing system and the high window temperatures achievable in supersonic flight make this a prime consideration. Window radiation can saturate the detector, and non-uniform radiation from the window can produce false signals which will mislead a tracking system, or produce ghosts in a reconnaissance system.
4. The window must not impose an unreasonable drag penalty on the vehicle.
5. The window must withstand the mechanical and thermal loads induced by high-speed flight.

Other factors such as ease of fabrication, light weight, ground handling loads must also be considered. Most of these requirements depend in whole or in part on the aerodynamic heating experienced in supersonic flight.

21.2.2. Window Heating

21.2.2.1. *Aerodynamic Heat Transfer.* The rate of aerodynamic heat transfer per unit area, q, of the wall (window) w can be written

$$q = h(T_r - T_w) \tag{21-1}$$

where h = heat-transfer coefficient
T_w = wall temperature
T_r = recovery temperature

The heat-transfer coefficient is a function of the wall geometry, the flight conditions, and to some extent the wall temperature. The recovery temperature is equal to the temperature of the wall when there is no aerodynamic heat transfer. This temperature is usually defined by means of a recovery factor r:

$$T_r = T(1 + 0.2rM^2) \tag{21-2}$$

If the wall (window) is assumed to be thermodynamically thin, *i.e.*,

$$\frac{b_w h}{k_w} << 1$$

and the radiation and conduction losses are negligible, the aerodynamic heat-transfer rate can be equated to the rate of enthalpy increase within the wall, or

$$h(T_r - T_w) = \rho_w b_w c_{p_w} \frac{dT_w}{dt}$$

In the last two expressions, b = window thickness, k = thermal conductivity, ρ = density, and t = time.

If there is a step change in the aerodynamic conditions, *e.g.*, velocity (thus a step change in h and T_r), the solution is

$$\frac{T_r - T}{T_r - T_i} = \exp - \left(\frac{h}{\rho_w c_{p_w} b_w} \right) t \tag{21-3}$$

where T_i = initial temperature at $t = 0$

At the higher Mach numbers (higher window temperatures) the radiation heat transfer plays a significant part. If this is to be considered, the heat-balance equation becomes

$$h(T_r - T_w) = \rho_w b_w c_{p_w} \frac{dT_w}{dt} + \sigma \epsilon_e(T) T_w^4 \tag{21-4}$$

where the equivalent total wall emissivity ϵ_e is a strong function of temperature because of its large spectral variation and σ = Stefan-Boltzmann constant. Equation (21-4) can be integrated step by step to yield a more accurate temperature history. The true equilibrium temperature ($dT/dt = 0$) can also be found from Eq. (21-4).

In a range of aerodynamic variables, it is more convenient to normalize the heat-transfer coefficient with respect to the local free-stream properties. The normalized heat-transfer coefficient is the Stanton number (St),

$$St = \frac{h}{\rho_e u_e c_p} \tag{21-5}$$

where u = velocity and e = flow properties of external edge of boundary layer.

For a more rigorous treatment of aerodynamic heat transfer, see [1] and [2]. Details of the temperature distribution within windows, given the aerodynamically imposed boundary conditions, can be developed by methods described in [3] and [4].

21.2.2.2. *Heat Transfer to Flat Plate.* The boundary-layer flow over a flat surface can be either laminar or turbulent. In most flight situations where the flat window is more than a foot or so from the nose or leading edge of the vehicle, the flow is turbulent. Transition from laminar to turbulent flow takes place at Reynolds numbers Re of approximately 10^5 to 5×10^6, the lower value being more likely in an actual flight vehicle with the usual surface roughness.

In the laminar case, the results of Van Driest [5] can be used directly. These extensive heat-transfer calculations are summarized in Fig. 21-1 as curves of $St\sqrt{Re}$ *versus* Mach number for various ratios of wall to free-stream temperature.

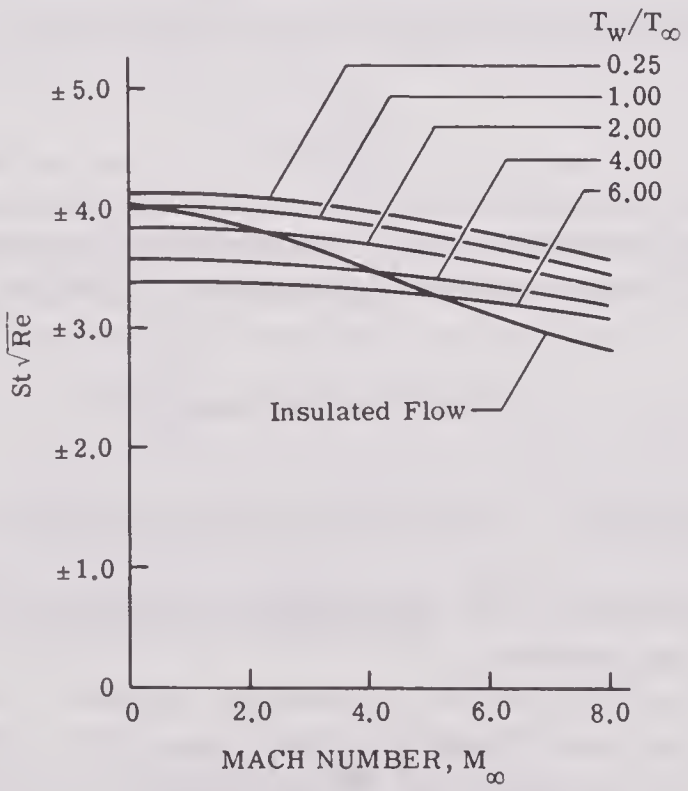


FIG. 21-1. Local heat-transfer coefficient for laminar boundary layer of a compressible fluid flowing along a flat plate, T_∞ = 400°R [5].

Experimentally, the recovery factor, r, is given for laminar flow as $\sqrt{Pr}$, or about 0.848 as an average over normal flight conditions (Pr = Prandtl number).

In the turbulent case, use is made of the Reynolds analogy, which relates the Stanton number to the local skin friction coefficient c_f through the equation

$$St = \frac{1}{s}\frac{c_f}{2} \tag{21-6}$$

Here $1/s$ is a measure of the accuracy of the analogy. For current purposes assume $s = 1$ although in some cases s can be as low as 0.8. Extensive measurements of skin friction have been made, and a semi-empirical relation which holds generally for incompressible turbulent flow has been found [6] which adequately predicts the local c_f:

$$c_{f_i} = 0.370/(\log_{10} Re_x)^{2.584} \tag{21-7}$$

where i = incompressible and x = some arbitrary distance.

Flow compressibility has a marked influence on the friction coefficient. To account for this, the ratio of compressible to incompressible friction coefficient for air is [1,6]

$$\frac{c_f}{c_{f_i}} = \left[0.5\left(1 + \frac{T_w}{T_e}\right) + 0.0394M^2\right]^{-0.648} \tag{21-8}$$

In the special case where the wall temperature is equal to the recovery temperature (insulated wall case)

$$\frac{c_f}{c_{f_i}} = \left[1 + 0.129M^2\right]^{-0.648} \tag{21-9}$$

The turbulent recovery factor, r, is approximately $(Pr)^{1/3}$, or about 0.896.

The insulated wall case serves as a handy reference for heat-transfer calculations. At $T_w = T_r$ there is, of course, no heat transfer, however, the Stanton number for the insulated case can be computed for these conditions. This corresponds rigorously to

$$St = \frac{dq}{dT}(\rho_e u_e c_p)^{-1}$$

at $T_w = T_t$ and can be used to determine the heat-transfer coefficients at wall temperatures in the vicinity of T_r. Figure 21-2 shows the variation of both laminar and turbulent Stanton numbers as a function of Mach and Reynolds numbers.

21.2.2.3. *Heat Transfer to a Hemisphere.* The aerodynamic flow to a hemispherical nose at the stagnation point (*i.e.,* the foremost point on the hemisphere, regardless of angle of attack), is laminar. As the air accelerates in passing from the stagnation point to the base of the hemisphere, it can go through transition and the boundary layer generally becomes turbulent. This happens at Reynolds numbers around 4×10^5.

For the laminar case, Sibulkin [7] has shown that locally the Stanton number on a hemisphere stagnation point can be written

$$St = 0.763\left(\frac{c_p\mu}{k}\right)^{-0.6}\left(\frac{\beta D}{u}\right)^{1/2}\left(\frac{\rho u_\infty D}{\mu}\right)^{1/2} \tag{21-10}$$

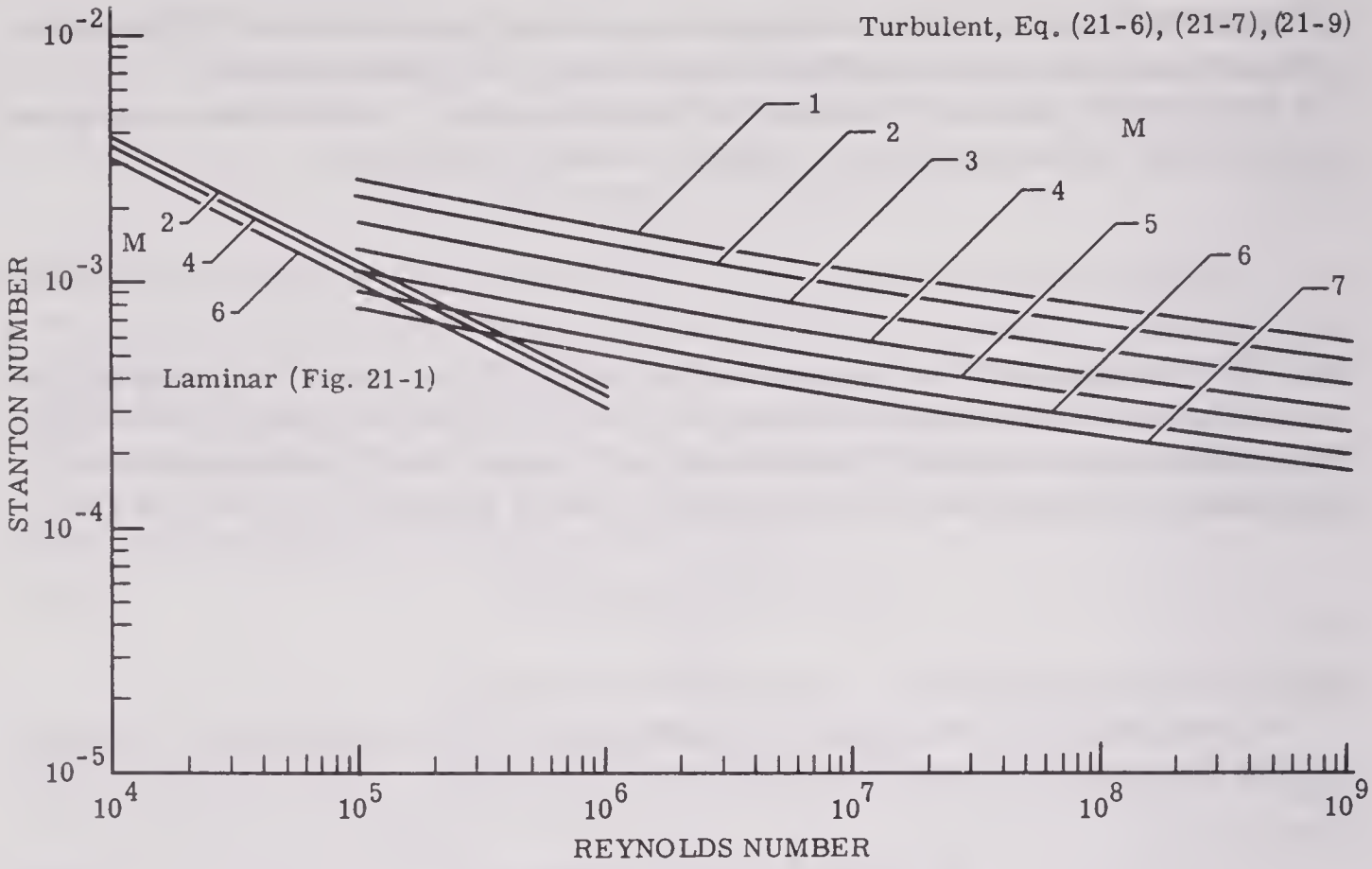


FIG. 21-2. Stanton number as a function of Reynolds number for an insulated plate.

where D = diameter, β = velocity gradient, μ = viscosity, and all of the fluid properties are evaluated at the local conditions just outside the boundary layer. In the case of a stagnation point (*s.p.*) this corresponds to stagnation temperature and pitot pressure. This reduces Eq. (21-10) into terms of free-stream flow:

$$St = 0.763 \left(\frac{\beta D}{u}\right)^{1/2} \left(\frac{\rho_\infty u D}{\mu_\infty}\right)^{-1/2} \left(\frac{c_p \mu}{k_\infty}\right)_{s.p.}^{-0.6} \left(\frac{\rho^0}{\rho_\infty}\right)^{1/2} \left(\frac{\mu^0}{\mu_\infty}\right)^{0.5} \tag{21-11}$$

where $(c_p\mu/k)_{s.p.}$ = stagnation Prandtl number (approximately 0.72 for normal flight conditions)

$\rho_\infty UD/\mu_\infty$ = free-stream Reynolds number based on hemisphere diameter

From [1],

$$\beta \frac{D}{U} = \sqrt{8 \frac{\rho_\infty}{\rho_{02}}}$$

If it is assumed that $\mu \propto T^{0.76}$ and the perfect gas equations of [8] are used, Eq. (21-11) then reduces to the general laminar Stanton number equation

$$St \sqrt{Re_D} \sqrt{U/D} = 9.37 \left(\frac{M^2}{M^2 + 5}\right)^{0.25} \left(1 + \frac{M^2}{5}\right)^{0.38} \tag{21-12}$$

This gives the heat-transfer rate at the stagnation point where the recovery factor, r, equals 1. The laminar heat-transfer rate decreases away from the stagnation point. Reference [9] gives methods for determining the laminar heat-transfer distribution over the surface of the hemisphere.

When the boundary layer is turbulent, the Stanton number has a maximum at about 40° from the stagnation point, then decreases further aft. To compute the local turbulent heat-transfer rate on the face of the hemisphere the "local flat plate" method can be used with good accuracy.

This approach calls for utilization of Eq. (21-6), (21-7), and (21-8) to determine the local Stanton number. All of the flow properties are evaluated at "local" conditions, that is, at the pressure, velocity and temperature just outside the boundary layer at any angular station θ. The Mach number distribution over the face of a hemisphere is insensitive to free-stream Mach number (Fig. 21-3). The local pressure is given by [8].

$$\frac{P_e}{P_{02}} = \left(1 + \frac{M_e^2}{5}\right)^{-7/2} \tag{21-13}$$

where

$$\frac{P_{02}}{P_\infty} = \left(\frac{6M_1^2}{5}\right)^{7/2} \left(\frac{6}{7M_1^2 - 1}\right)^{5/2}$$

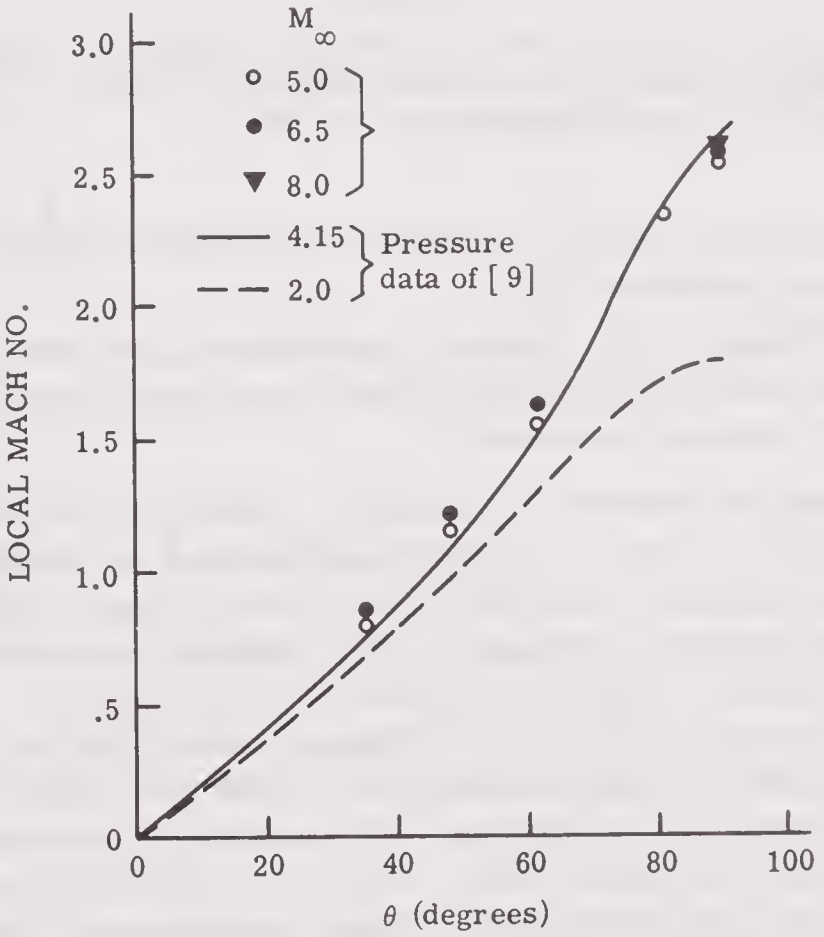


FIG. 21-3. Local Mach number on hemisphere.

The local static temperature

$$\frac{T_e}{T_{0_\infty}} = \left(1 + \frac{M_e^2}{5}\right)^{-1} \tag{21-14}$$

since the stagnation temperature is unchanged by a shock wave.

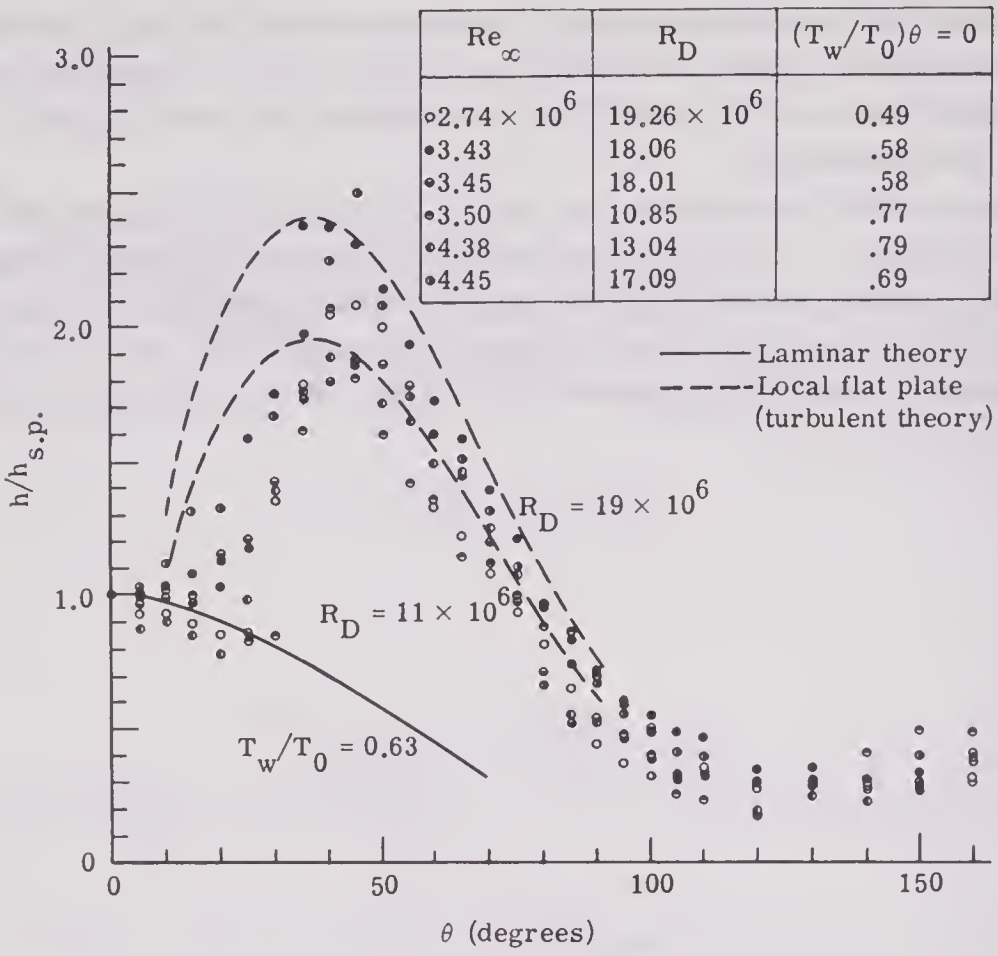


FIG. 21-4. Variation with θ of the ratio of local heat-transfer coefficient to the value at the stagnation point at $M_\infty = 2.0$ and several free-stream Reynolds numbers [10].

A typical case is shown in Fig. 21-4, where both theoretical and experimental heat-transfer distributions are shown [10].

21.2.3. Typical Results. The following subsections apply some of the basic equations previously discussed to typical flight situations to determine what sorts of temperatures and heating rates can be expected.

21.2.3.1. *Maximum or Recovery Temperature.* If there are no conduction losses and the window emissivity is zero, the recovery temperature can be determined directly from Eq. (21-2) by supplying the appropriate recovery factor, 0.848, for laminar flow and 0.896 for turbulent flow. However, if the radiation heat transfer cannot be neglected, Eq. (21-4) must be solved, with $dT_w/dt = 0$ at equilibrium.

The actual total emissivity of the window depends strongly on the particular material selected, and on the instantaneous temperature. However, here ϵ_w will be considered constant. A total emissivity of 0.25, over all wavelengths, is characteristic of some materials over a large temperature range. Figure 21-5 shows the maximum or equilibrium temperatures at three altitudes and at Mach numbers up to 7 for the flat plate, and for a hemispherical window 6 in. in diameter.

21.2.3.2. *Rate of Temperature Increase.* To give an indication of the relative rate of temperature increase, Eq. (21-3) is used. The equation for a time constant, τ, is

$$\tau = \frac{\rho_w c_{p_w} b_w}{h} \tag{21-15}$$

and gives the time to reach

$$\frac{T_r - T}{T_r - T_i} = \frac{1}{e} \approx 0.368 \tag{21-16}$$

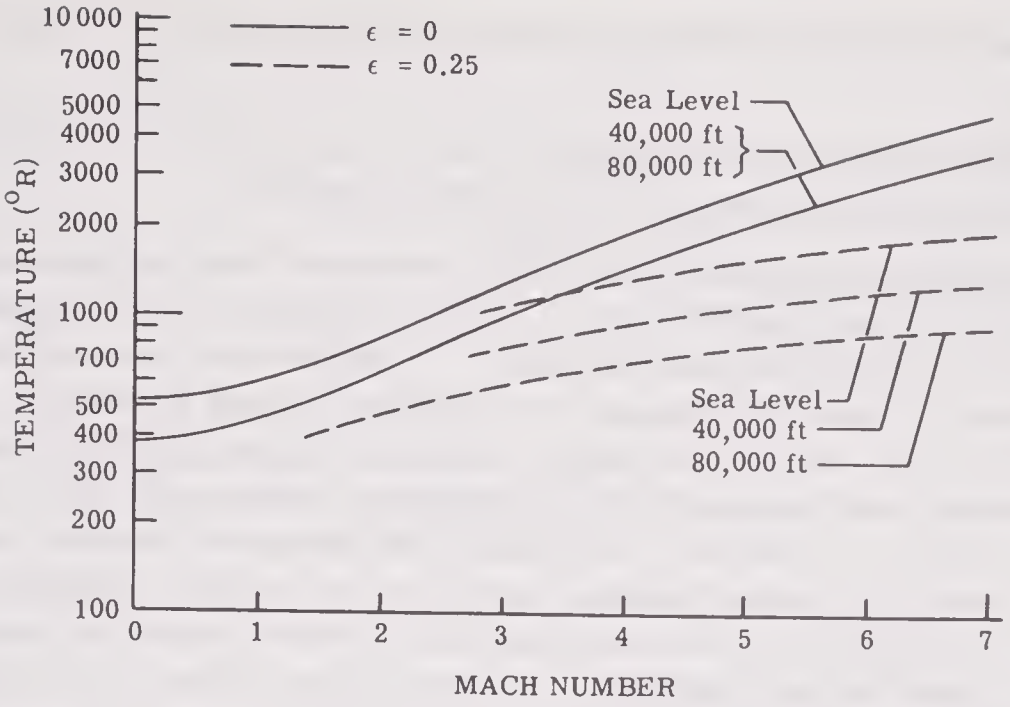


FIG. 21-5. Maximum temperature due to aerodynamic heating, turbulent flat plate, $x = 1$ ft.

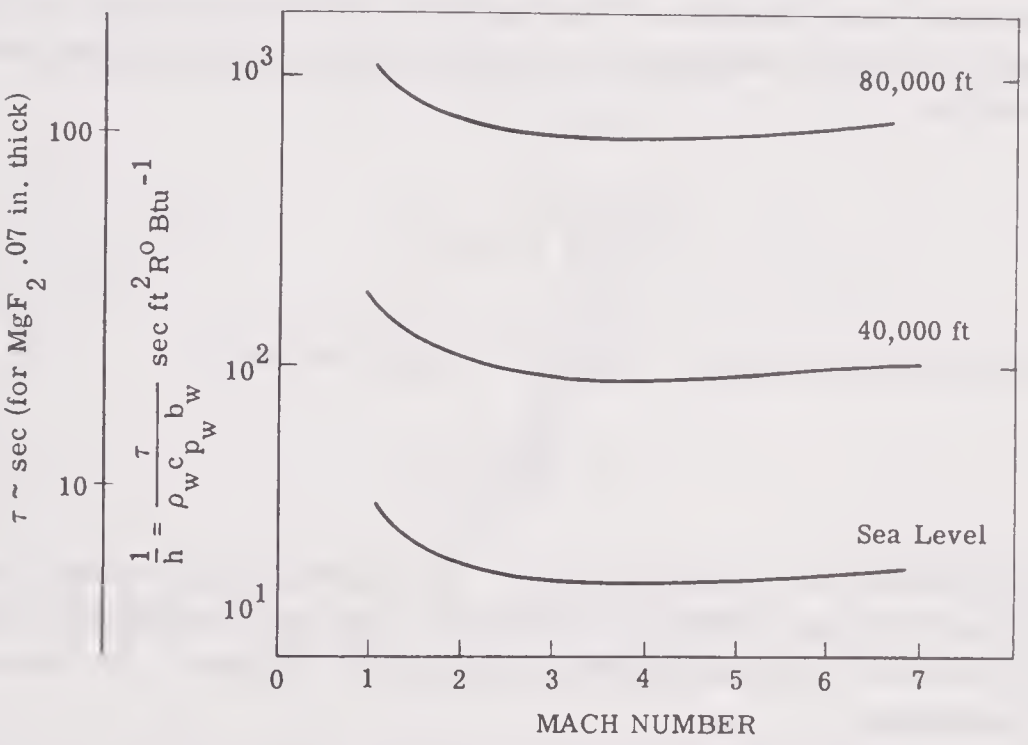


FIG. 21-6. Time constant for aerodynamic heating.

in the absence of radiation. In Fig. 21-6 the generalized time constant is given for the conditions of Fig. 21-5, and a typical material, MgF_2, 0.07 inch thick, is used to get a specific time constant.

21.2.4. Window Transmission and Radiation at Elevated Temperatures. Aerodynamic heating can produce changes in the infrared characteristics of window materials. A major problem at elevated temperatures is a loss in transparency, particularly at the longer wavelengths. Transmission characteristics of materials at high temperatures are covered in Chapter 8. Any body, semi-transparent or opaque, will radiate thermal energy in ever-increasing amounts as the temperature rises. Two different types of effects are produced by a radiating window:

(1) Because of strong temperature gradients, or local "hot spots," false targets can appear.

(2) A large irradiance on the infrared detector, even though uniformly distributed, can saturate the detector and decrease its sensitivity.

The susceptibility of a particular guidance system to false targets associated with nonuniform temperature distributions is greatly dependent on the details of the discrimination and tracking methods employed. A window which is nearly hemispherical does not generally have large temperature gradients along its surface, even though the temperature level can be quite high. A pointed window, such as a cone, can have very high temperature gradients along its surface, particularly when used at high speeds and moderate altitudes. Any small protuberances optically distort the incoming radiation and produce local hot areas on the window. Generally, these deviations are spread out by the thermal conductivity of the window and are close enough to the optical system to be out of focus. A point on the window is not likely to appear as an exact replica of a target, but will produce additional noise in the tracking system, thereby effectively reducing its sensitivity.

The problem of saturation of the infrared detector by the window radiation is usually more critical than the "false target problem" and can be treated more precisely. Details of the particular optical techniques employed do not influence the saturation calculation, and a general analysis is possible. If the field of view is assumed to be small, and the window is assumed to be isothermal and fills the entire field of view, the radiance B can be expressed as

$$B = \frac{1}{\pi} \int w_{\lambda_{bb}} \epsilon \, d\lambda \tag{21-17}$$

Now the effective radiant power P incident on the collector is

$$P = BA_c\pi(\Omega)^2 \tag{21-18}$$

where Ω is the half-angle of the field of view. By neglecting the losses in radiant power due to any space filtering, refractive optical elements, etc., the effective radiant power falling on the detector is equal to that falling on the collector. Therefore, the irradiance-at the detector is

$$H_d = \frac{P}{A_d} = B\pi\Omega^2 \frac{A_c}{A_d} \tag{21-19}$$

By geometry

$$\frac{A_c}{A_d} = \frac{1}{4}\left(\frac{D}{f}\right)^2 \Omega^2$$

where f = focal length. The irradiance at the detector is a function of only the window radiation and the focal length-to-diameter ratio of the collector:

$$H_d = B\frac{\pi}{4}\left(\frac{f}{D}\right)^{-2} \tag{21-20}$$

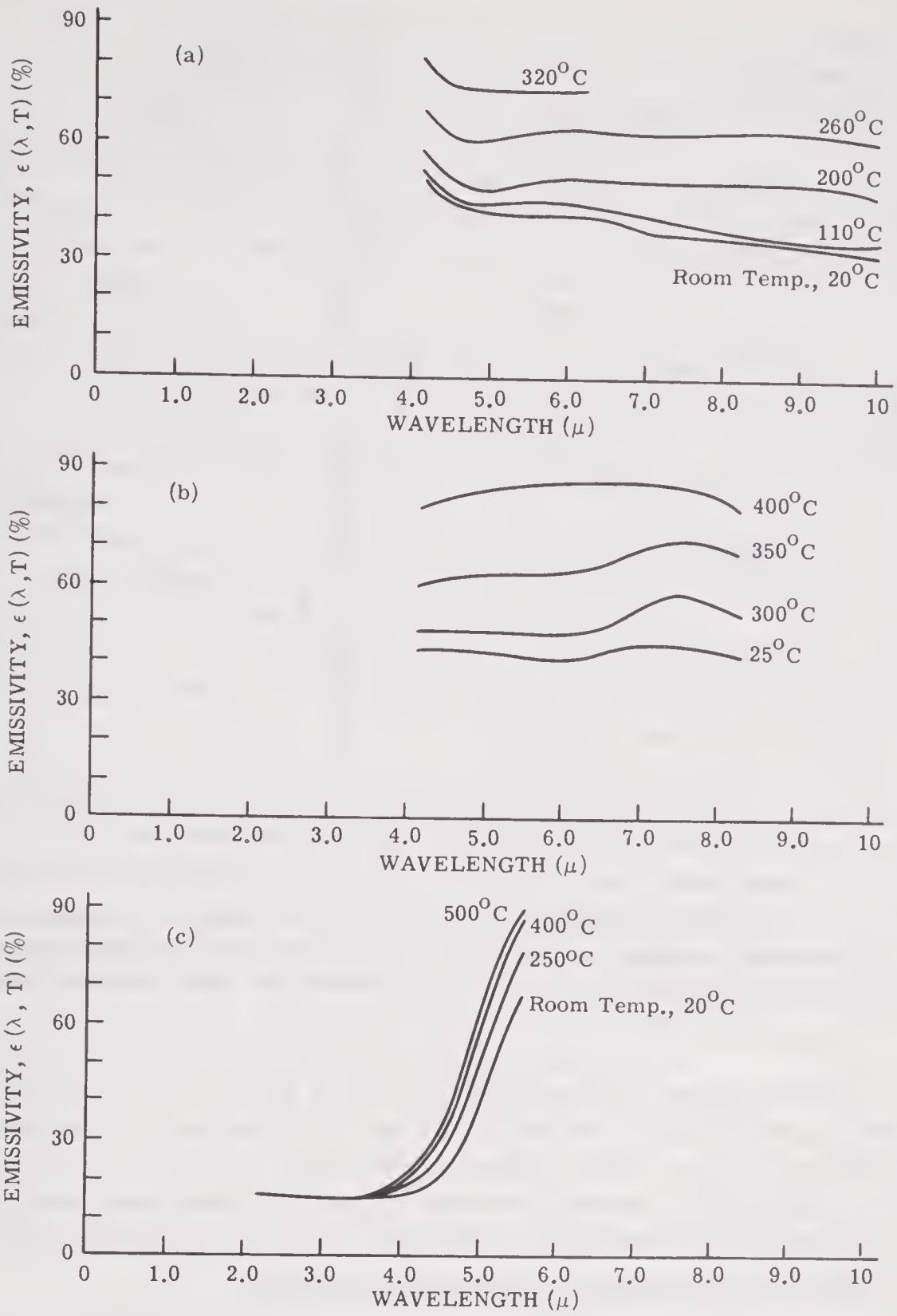


FIG. 21-7. Emissivity of (*a*) germanium (single crystal, 1.14 mm thick); (*b*) silicon (single crystal, 4.16 mm thick); (*c*) synthetic sapphire (4.11 mm thick) [11].

The emissivities of several window materials are shown in Fig. 21-7, others are covered in Chapter 8.

Figure 21-8 shows the detector effective spectral irradiance due to a typical window of MgF_2 at various temperatures.

In a conventional air-to-air missile application, the window temperature can reach several hundred degrees. Figure 21-8 shows that at the detector the background irradiance caused by the window is many orders of magnitude greater than the irradiance from the target. The saturation effects of elevated window temperature are shown in Fig. 21-9 as the relative magnitude of the minimum detectable signal (MDS), normalized for MDS at a window temperature of 50° C, as a function of window temperature. It is assumed that no window radiation outside the 3-μ to 5-μ region falls on the

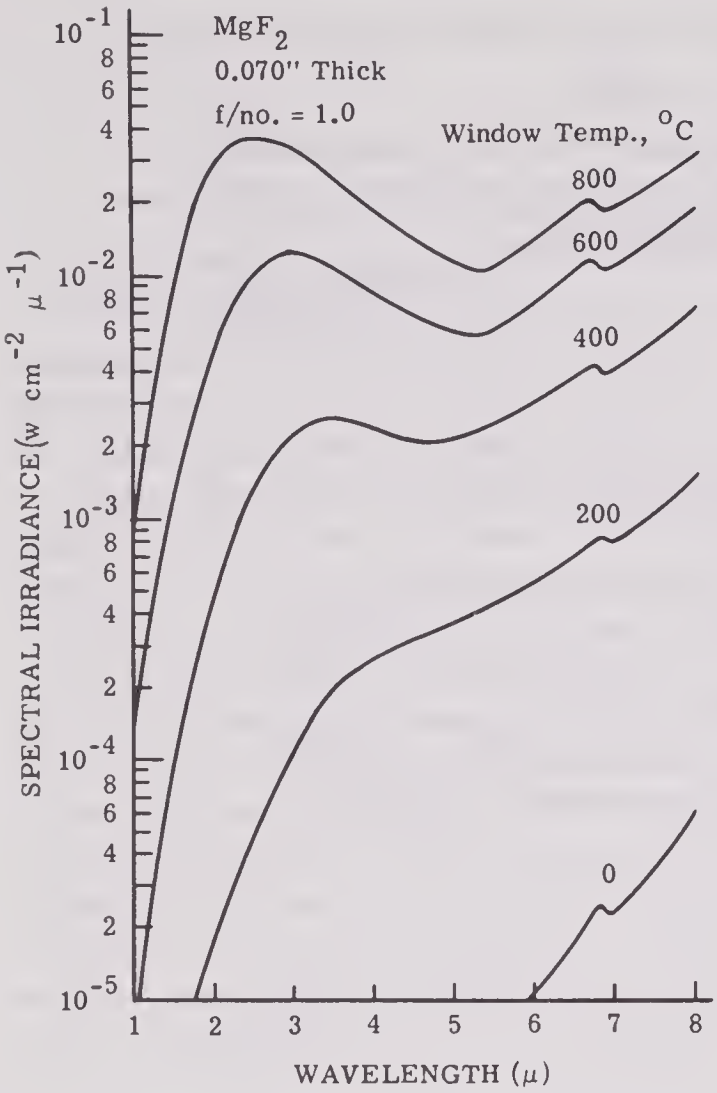


FIG. 21-8. Effective spectral irradiance at a detector with an MgF_2 window *vs.* wavelength, at various temperatures.

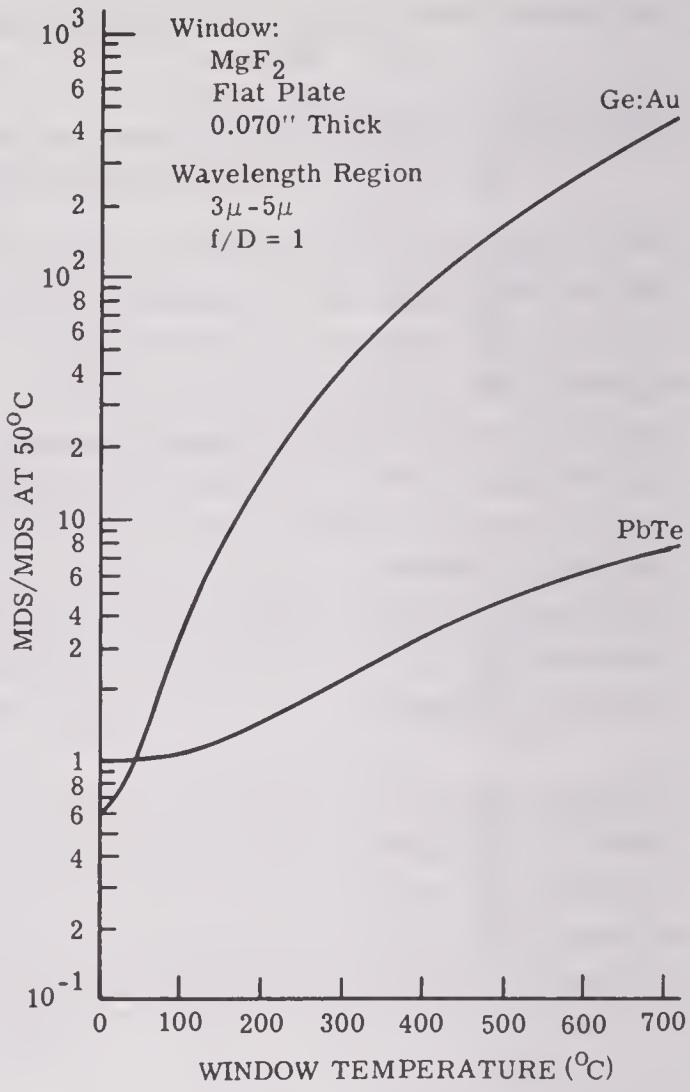


FIG. 21-9. Relative minimum detectable signal for Ge:Au and PbTe detectors as a function of MgF_2 window temperature.

detector. A PbTe detector can lose sensitivity by a factor of 8 if the MgF_2 window temperature is raised from 50° C to 700° C, and the Au:Ge detector can lose sensitivity by a factor of 360 for the same window temperature change. It should be noted that MgF_2 is a good infrared window material between 3 μ and 5 μ, and other materials would probably produce more of a saturation effect over the same temperature range.

21.2.5. Methods of Alleviating Hot-Window Problems

21.2.5.1. *Optical-System Parameters.* Several optical-system parameters have a considerable influence on the overall performance through the effects of window radiation. They include:

(1) Spectral region of sensitivity—the spectral distribution of the window radiation can alter the optimum wavelength limits from those indicated by target-detector-background considerations alone.

(2) Detector type—of the several detectors that can operate in the desired spectral region, improved saturation characteristics can produce a greater signal-to-noise ratio than slightly greater detectivity in the presence of strong window radiation.

(3) Field of view—reduction of the field of view is important since the window irradiation on the cell and the cell area are reduced.

(4) Window location—for systems not requiring a nose installation, careful selection of the window location so as to avoid regions of high heat transfer (shock-wave impingement, boundary-layer reattachment, etc.) can reduce window radiation.

21.2.5.2. *Delay in Temperature Rise of the Window.* In many systems, the time of operation is short enough and/or the closing rate on the target is rapid enough (the target signal increases) so that a simple delay in the temperature rise of the window constitutes a satisfactory solution to hot-window problems.

Several delaying methods can be used. The most obvious approach is to precool the window. When the tactical situation permits (*i.e.,* internal missile storage in flight, ground launch) a marginal condition can be made operational without complicating the basic system design.

Assuming that the window material has the highest value of ρc_p consistent with the optical requirements, the time constant can be lengthened by increasing the window thickness, b. This is fairly effective with window materials in which the transmission loss is due primarily to reflection rather than absorption. Optical distortion and window weight also limit this approach. Again, a marginal condition can be improved by this method.

Another solution involves reducing the heat-transfer coefficient. A flat window mounted along the side of a vehicle can be recessed below the vehicle surface, thereby separating the boundary layer from the window surface.

Assuming that the cavity depth is of the same scale as the thickness of the boundary layer and that the boundary layer is always turbulent, two separate flow configurations can exist over such a shallow cavity. The first, "closed cavity," flow exists when the boundary layer, initially separated as it crosses the upstream edge of the cavity, attaches to the cavity floor and then separates again ahead of the downstream edge of the cavity (Fig. 21-10*a*). The second, "open cavity," flow exists when the boundary layer remains separated over the entire length of the cavity (Fig. 21-10*b*). The flow configuration goes from closed to open as the cavity length-to-depth ratio is decreased. As the boundary layer becomes thicker with respect to the cavity depth, the transition from closed flow to open flow becomes more gradual and the flow configuration is not as sensitive to changes in length-to-depth ratio, L/H. Roughly, for a uniform depth cavity, length-to-depth ratios less than about 10 will produce open-cavity flow.

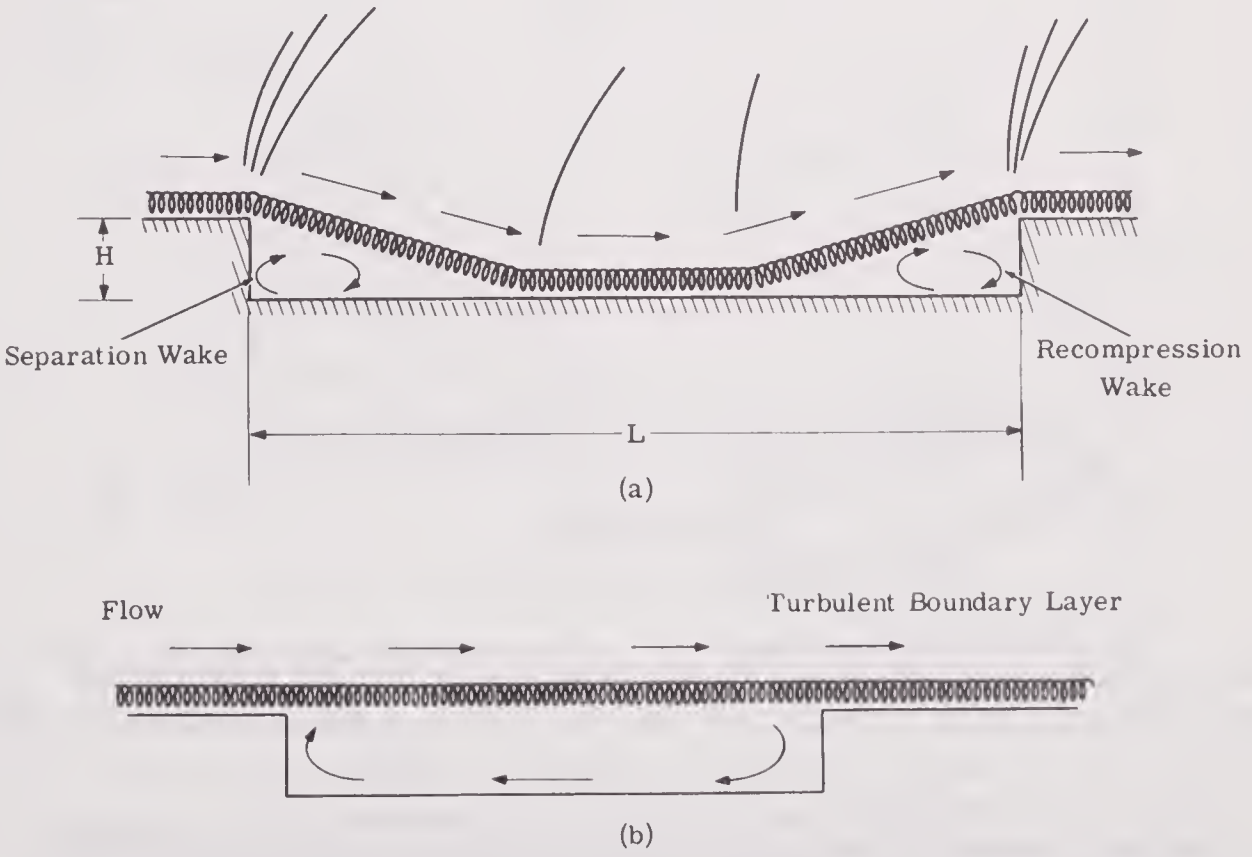


FIG. 21-10. Flow configurations. (*a*) Closed cavity; (*b*) open cavity.

Theoretical treatment of aerodynamic heat transfer in separated flow has not proved too satisfactory because of lack of available experimental data. For experimental data measured in a configuration similar to the one under consideration at moderate Mach numbers, see [12]. The heat-transfer coefficient, h, on the cavity floor, normalized with respect to the heat-transfer coefficient for an attached flow at the same body station and Reynolds number, is shown in Fig. 21-11 and 21-12.

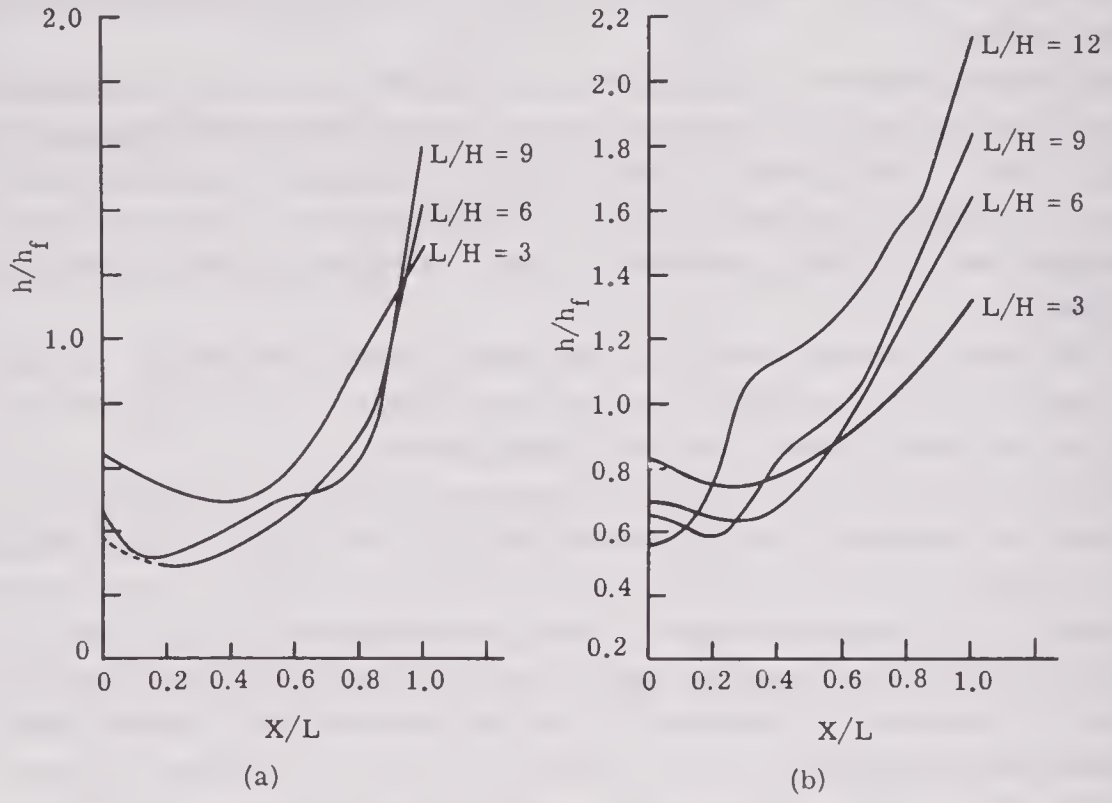


FIG. 21-11. Variation of the heat-transfer distribution on the floor of the cutout as a function of L/H at $M_0 = 2.9$ [12]. (*a*) Thin oncoming boundary layer; (*b*) thick oncoming boundary layer.

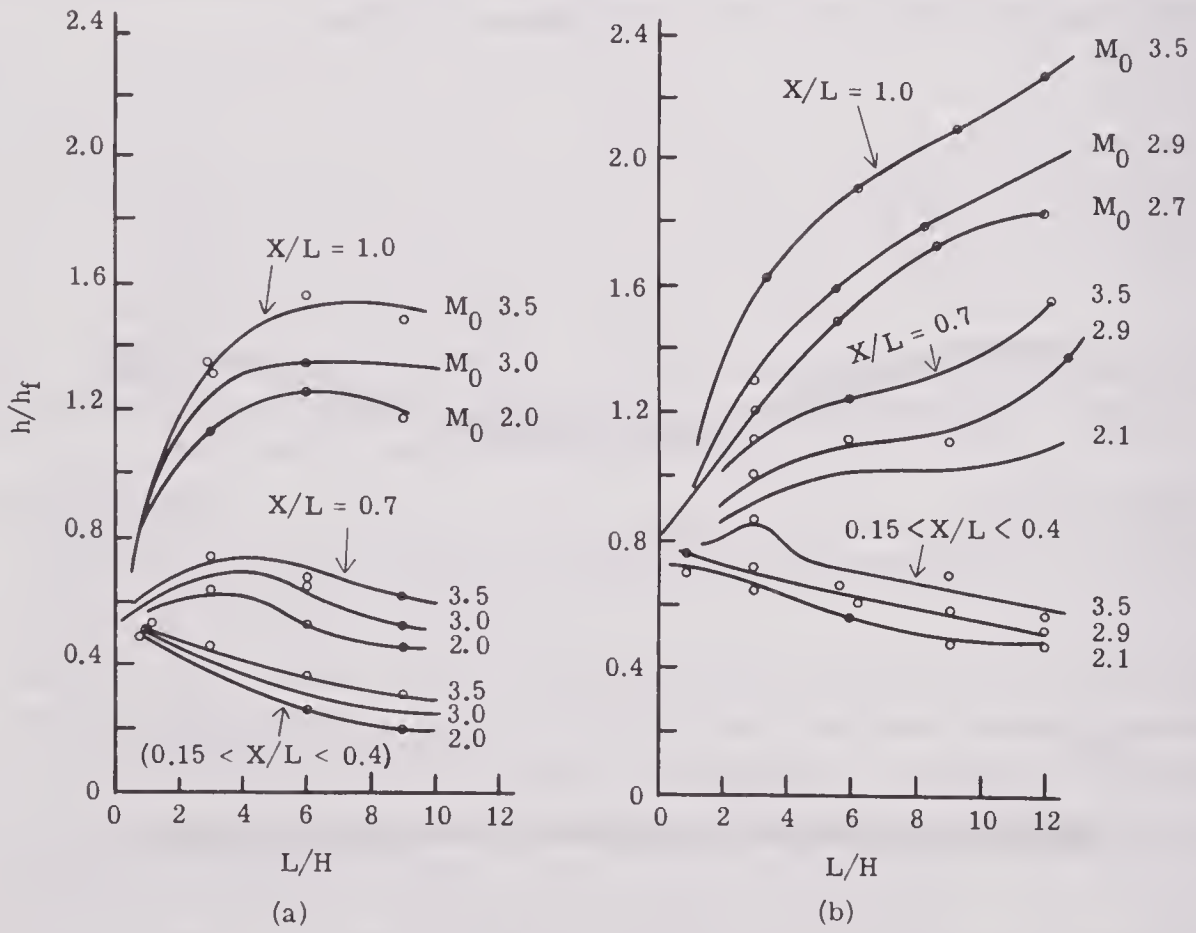


FIG. 21-12. Parametric map of heat-transfer coefficient ratio (minimum, maximum, and at $X/L = 0.7$) as functions of L/H and M_0, (*a*) thin boundary layer; (*b*) thick boundary layer [12]. Note: points are averages over all tests.

For all values of L/H the heat-transfer-coefficient ratio drops considerably below 1 in the separation wake, and rises above 1 in the recompression wake, reaching a maximum at the downstream edge of the cavity. The typical h/h_f distributions plotted in Fig. 21-11 show that an increase in the boundary-layer thickness does not change the general character of the curves, but does raise significantly the average level of heat transfer. In Fig. 21-12 it can be seen that the relative heat-transfer coefficient decreases with decreasing Mach number and a higher ratio of L/H produces a greater maximum heat-transfer coefficient.

21.3. Refraction by the Field of Flow

21.3.1. Index of Refraction of Air. In the analysis of aerodynamic influences on electromagnetic radiation, the index of refraction n, of air must be considered. The major change in index of refraction of air takes place with a change in density, ρ. The variation of index of refraction for transparent substances as a function of density is [13]

$$\frac{n^2 - 1}{n^2 + 2} = k\rho \tag{21-21}$$

This is the Lorenz-Lorenz law which for gases ($n \approx 1$) reduces to the simpler Gladstone-Dale law:

$$n - 1 = k\rho \tag{21-22}$$

with virtually no loss in accuracy. Figure 21-13 shows the variation of k in Eq. (21-22) as a function of wavelength, as deduced from the data of [14]. Variations in index of refraction due to temperature effects on molecular structure and changes in atmospheric composition are neglected.

Figure 21-13 shows that the effects of refraction through nonuniformities in air are more pronounced in the visible portion of the spectrum than in the infrared. In the wavelength region beyond the near infrared ($\lambda > 1\ \mu$) the refractive index of air is nearly invariant with wavelength; thus one finds no dispersion and no chromatic aberration.

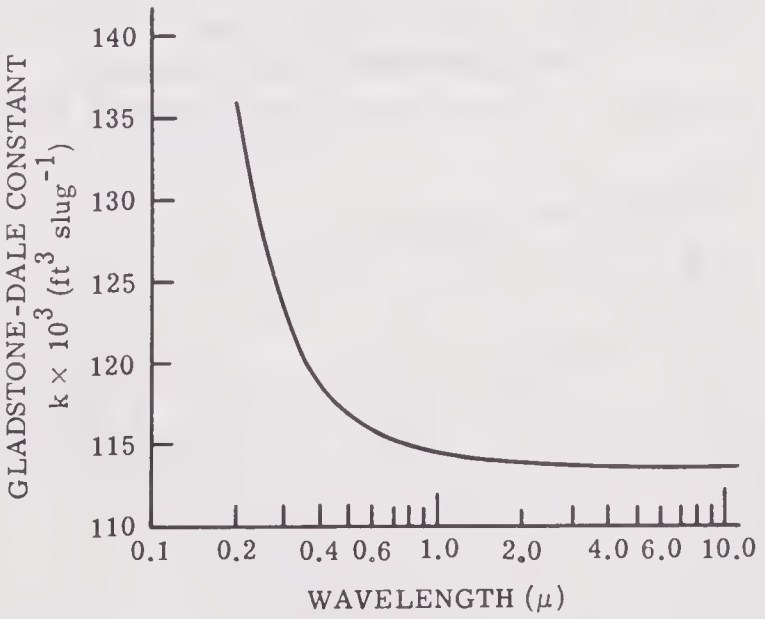


FIG. 21-13. Gladstone-Dale constant for air [14].

21.3.2. Shock-Wave Effects. As a body travels through the atmosphere, the air must adjust to accommodate it. At supersonic velocities the initial adjustment takes place across shock waves, which are discontinuities in the air properties.* Surfaces

*Shock waves are not exactly discontinuities but do have a finite thickness. This thickness is about five molecular mean free paths, which allows translational and rotational equilibrium of the molecules to be established. At sea level the shock wave is about 1 μ thick.

with sharp leading edges at moderate, positive angles of attack have these shock waves attached to the leading edge; relatively blunt bodies have shocks detached and standing some distance ahead of the body. The strength and orientation of these shock waves are determined by the flight Mach number and the geometry of the body.

The relation between the density ahead of the shock, ρ_1, and the density behind the shock, ρ_2, is

$$\frac{\rho_2}{\rho_1} = \frac{6M_\infty^2 \sin^2 \theta}{M_\infty^2 \sin^2 \theta + 2} \tag{21-23}$$

Equation 21-23 shows that the density ratio (shock strength) increases with the shock angle, θ, and the free-stream Mach number, M, approaching the maximum value of 6 for a normal shock at M_∞.

No convenient, explicit relations exist for determining the shock angle. However, curves of shock angle as a function of Mach number and flow inclination are given in [8]. Figures 21-14 and 21-15 show representative curves of the functions for both wedge and conical flow.

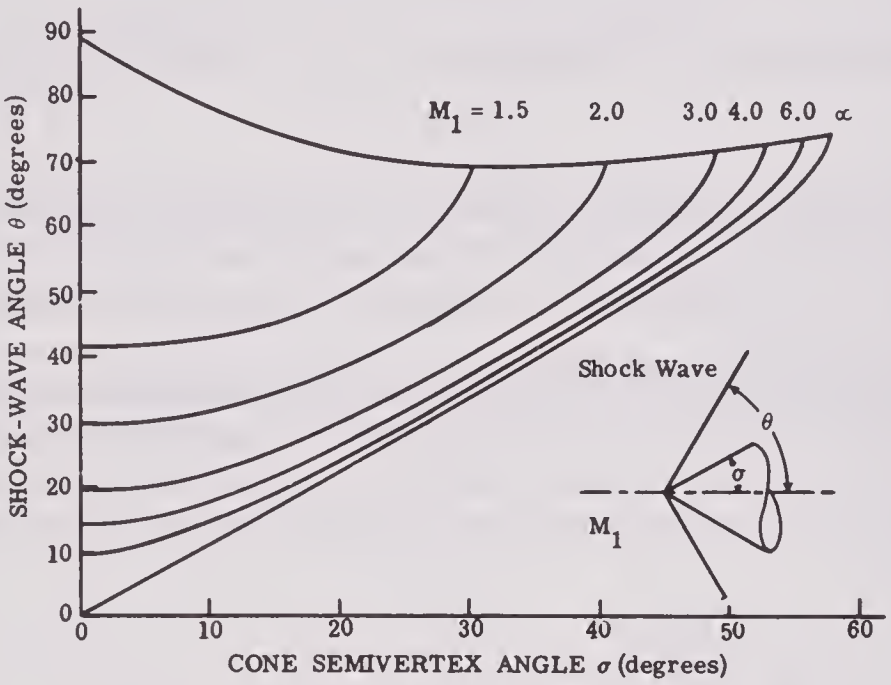


FIG. 21-14. Shock-wave angle for conical flow.

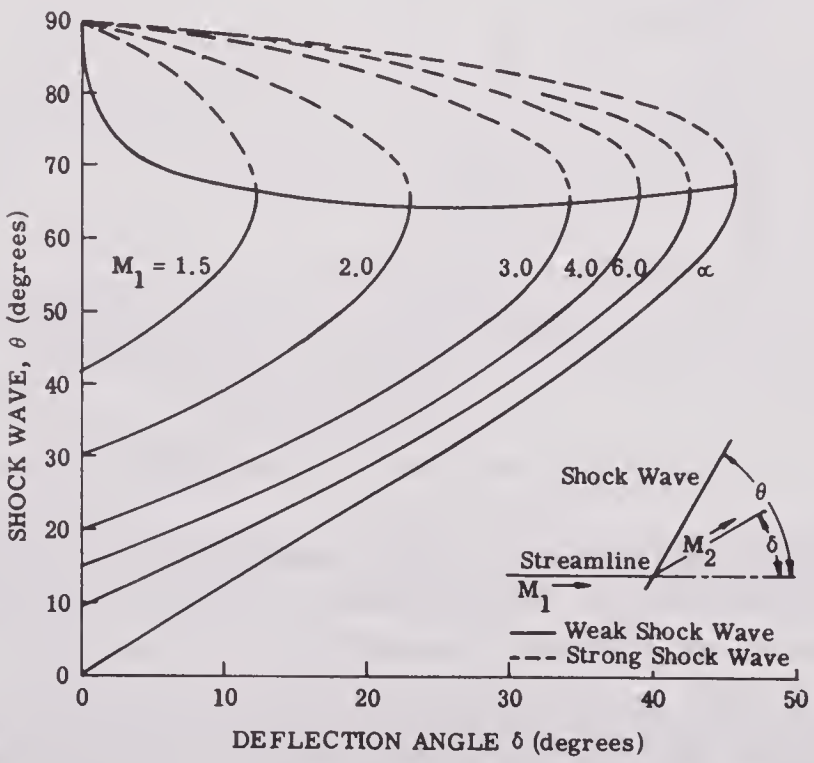


FIG. 21-15. Shock-wave angle for wedge flow.

A shock wave produced by solid surfaces is exceptionally steady and, because of the density change across it, is a good refracting surface. Since a plane shock wave is not necessarily perpendicular to the optical axis or parallel to the window surface, the effects of refraction must be taken into account even in the simplest case. Using Snell's law, the angle β_1 of the incident ray and the angle β_2 of the refracted ray, both measured from the normal to the shock plane, can be related by Eq. (21-24):

$$\frac{\sin \beta_1}{\sin \beta_2} = \frac{1 + k\rho_2}{1 + k\rho_1} \tag{21-24}$$

or alternatively,

$$\frac{\sin \beta_1}{\sin \beta_2} = \frac{1}{n_1} + \frac{n_1 - 1}{n_1} \frac{\rho_2}{\rho_1} \tag{21-25}$$

where n_1 is a special index value in the first medium. From Eq. (21-25), the angular deviation of a ray passing through a plane shock wave can be calculated as a function in incidence angle. This is shown in Fig. 21-16 for various density ratios and two altitudes. Figure 21-16 shows that the density ratio of 6 (the limiting value) at sea level is the worst possible refraction error. The angular deviation in this limiting case is seen to be about 1.5 mrad (300 sec of arc) at a 45° incidence angle. At the higher altitude, 50,000 ft, the maximum refraction error at 45° incidence is only 0.25 mil (50 sec), or about one-sixth of the sea-level value. The minimum density ratio shown is 2, which corresponds to a normal shock wave at about Mach 1.6. The deviation of the rays is only about one-fifth of the limiting value.

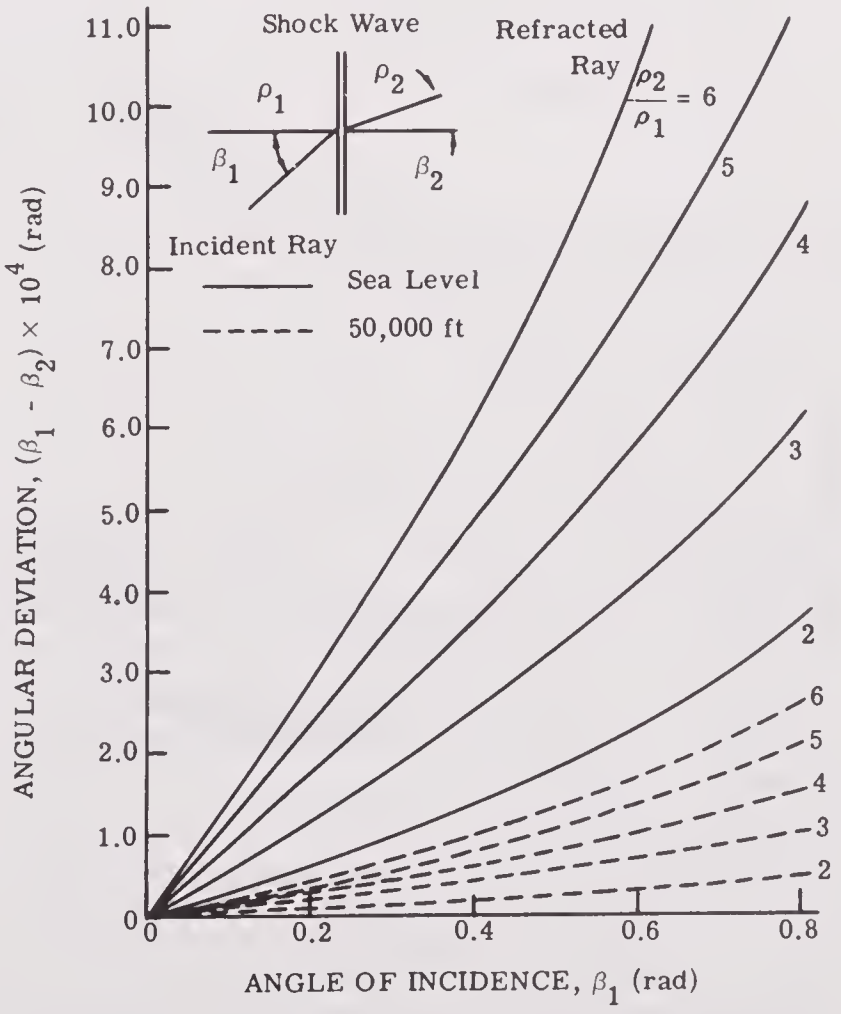


FIG. 21-16. Angular deviation of infrared radiation passing through plane shock waves.

Another way of considering the variation of plane shock-wave refraction effects is to fix the body geometry and allow the Mach number to change, producing variations in density ratio and shock angle. Selecting a flow inclination of 10°,

$$\frac{\cos(\theta + \phi_1)}{\cos(\theta + \phi_2)} = \frac{1 + k\rho_2}{1 + k\rho_1} \tag{21-26}$$

Figure 21-17 shows the angular deviation of a ray as a function of its angle with respect to the direction of flight. This corresponds directly to the error between the true and apparent target position as seen by an optical system situated in the wedge shown in the figure. From Fig. 21-17 it can be seen that a target along the direction of flight will appear 0.2 mil (42 sec) too high at Mach 2, and about 1 mil too high at Mach 5. In the limiting case of a Mach number of infinity, the error along the flight direction is 3.2 mils. This error is less as the incident ray becomes more nearly perpendicular to the shock wave, and when it is exactly perpendicular the error goes to zero.

Because of this shock wave, wide-angle photography suffers a distortion in the flight direction, even if the optical axis is perpendicular to it. If this distortion is evaluated as the difference in angular deviation of the rays at each edge of the field of view, the minimum distortion occurs when $\theta + \phi_1 = \pi/2$. At that point the slopes of the curves of Fig. 21-17 yield a distortion of about 0.3 mil per radian field of view at Mach 2, and about 0.7 mil per radian field of view at Mach 5. More explicitly, for example, a 1-rad field-of-view instrument, in a 10° wedge flying at Mach 5, would show two objects, one at each limit of the field of view, too close together by at least 3.6 ft for every mile of range.

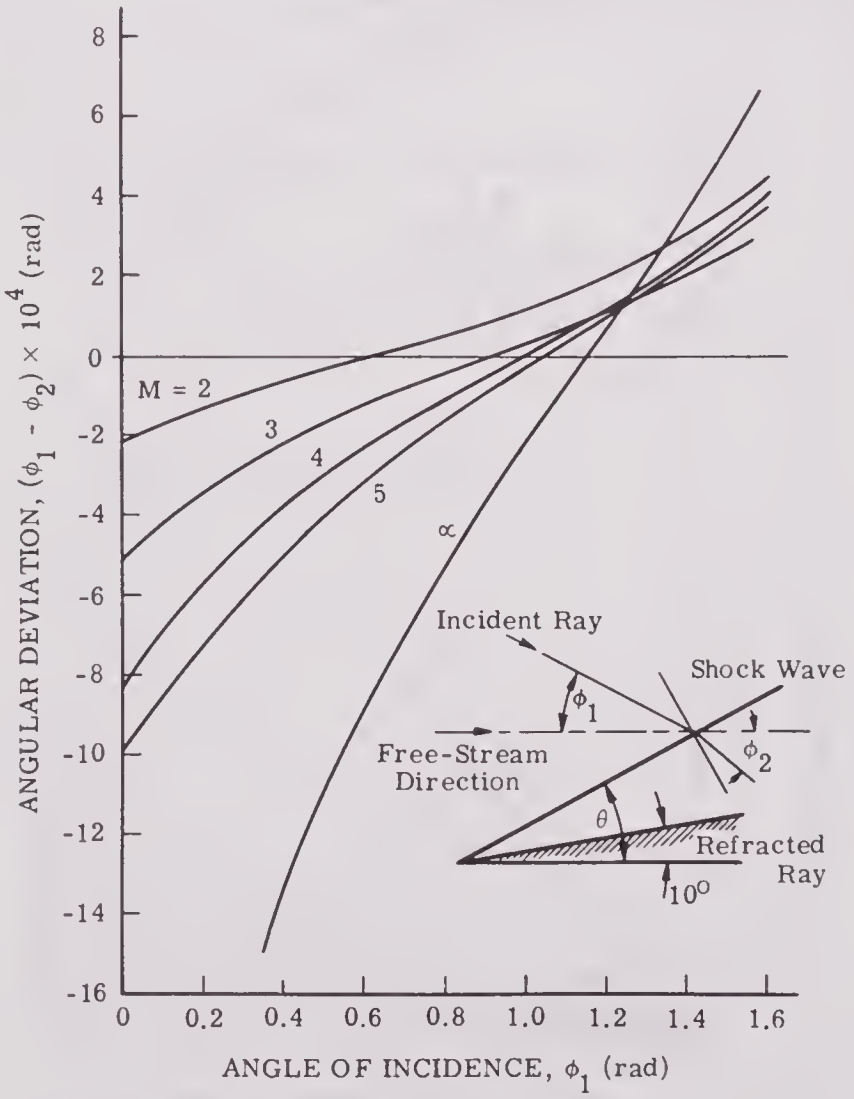


FIG. 21-17. Angular deviation of infrared radiation passing through plane shock waves produced by a 10° wedge (at sea level).

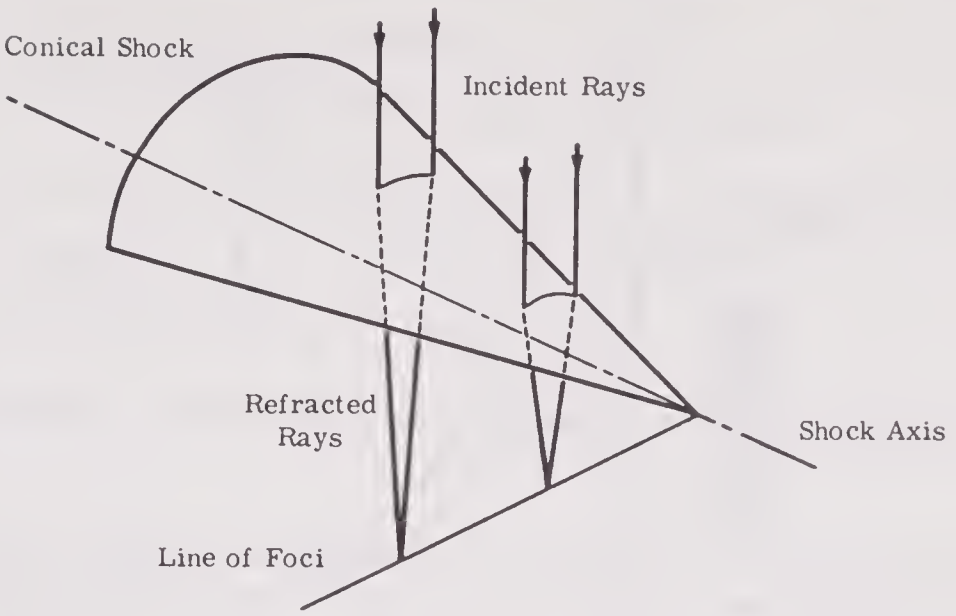


FIG. 21-18. Conical shock wave refraction of incident radiation.

The curved shock wave occurs when the shock is detached from the leading edge or when the body itself presents a nonplane surface to the flow (*e.g.,* curved cross section, or body width not large compared to its thickness). The introduction of curvature into the density discontinuity produces a focusing effect which, except in special cases, is astigmatic. To demonstrate the nature of the problem, consider first the conical shock-wave influence on an incident plane wave of radiation approaching perpendicular to the cone axis, as shown in Fig. 21-18. This curved shock acts as a lens and tends to focus incoming rays of parallel light. For the particular case of Fig. 21-18, the focal length is given by

$$\frac{f}{x} = \frac{n_1 \tan \theta}{n_2 - n_1} \tag{21-27}$$

where x = axial distance from the apex of the shock cone
θ = the shock angle.
n_2 = index value in second medium.

The locus of the "focal line" has been calculated for two cone half-angles, 10° and 20°, and the variation with Mach number is shown in Fig. 21-19. It can be seen that the shock lens becomes stronger (f/x less) with increasing cone angle and Mach number because of the increase in shock strength. The effect of this conical lens is to skew the focal plane of a normal optical system whose axis is aligned with the incoming radiation. In order to estimate the importance of this effect, consider an $f/1$ optical system with a 15-cm aperture, 40 cm from the apex of a 10° half-angle cone. If the source is a line parallel to the cone axis and at an infinite distance, the change in the image at the focal plane can be calculated only on the basis of geometrical changes (*i.e.,* the change in convergence angle of the rays). The magnitude of the defocusing due to the shock then can be compared to the diffraction limit of the simple optical system. Table 21-1 shows that the image defocusing due to the conical shock wave at Mach 3, in an uncorrected system, would be about five times as large as the theoretical resolution limit, and about ten times larger at Mach 5. Simply refocusing the system would bring the defocusing to within a factor of 2 of the theoretical resolution limit even at Mach 5. As the optical axis tilts toward the shock apex, the curvature normal to the incoming wave front becomes elliptical and then parabolic rather than circular, introducing further astigmatism. Also, as the optical path passes through regions near the apex of the shock cone, the radius of curvature of the shock decreases, producing a lens of greater power.

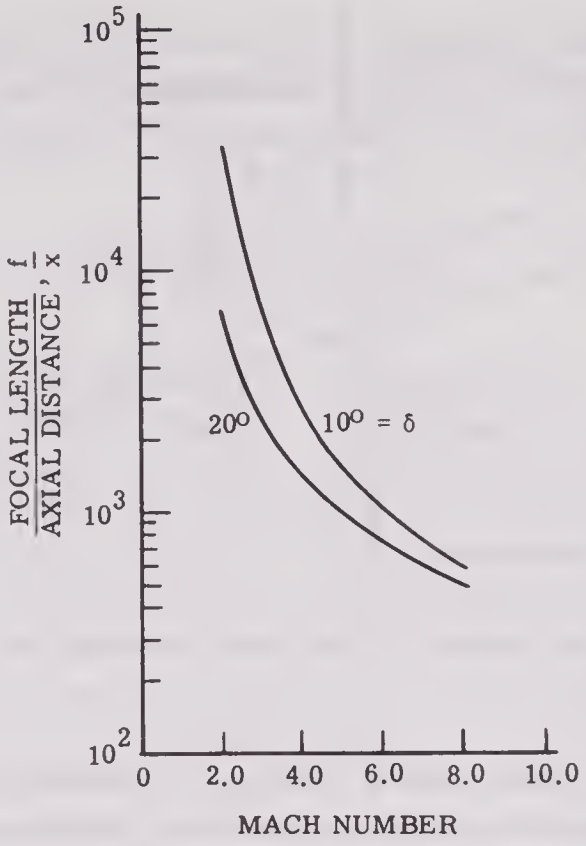


FIG. 21-19. Ratio of focal length to axial distance from apex for a conical shock wave.

TABLE 21-1. EFFECT OF CONICAL SHOCK WAVE ON *f*/1, 15-CM OPTICAL SYSTEM

		Mach = 3.0	*Mach* = 5.0
Theoretical resolution limit		2.4×10^{-5} rad	2.4×10^{-5} rad
Change (in focal length) due to shock	(fwd)	2.4×10^{-3} cm	6.4×10^{-3} cm
	(aft)	1.7×10^{-3} cm	4.6×10^{-3} cm
Angle subtended at image by 15-cm focus with shock	(fwd)	1.6×10^{-4} rad	4.3×10^{-4} rad
	(aft)	1.1×10^{-4} rad	3.1×10^{-4} rad
Maximum image angle subtended at optimum focus with shock		2.1×10^{-5} rad	5.8×10^{-5} rad

A more complicated optical mechanism exists in the case of flow about a conical body. This arises because the density of the air is not constant in the flow field between the shock wave and the body surface. To accurately compute this type of effect involves a complicated integration process; the approximate solutions are given in [15]. Typically, for a cone half-angle of 20° the contributions of the density gradients behind the shock wave are approximately 26% and 14% of the total error at Mach numbers of 2.0 and 3.5, respectively, when the ray is approximately perpendicular to the cone axis.

In the case of a hemisphere traveling at supersonic speeds, a nearly spherical detached shock forms ahead of the body and becomes part of the optical path. Assume, for example, a plane wave of radiation falling on the hemispherical shock along its axis of symmetry (Fig. 21-20). Rays striking an oblique portion of the shock surface should be brought to some sort of focus on the axis of symmetry. However, the change in index of refraction across the shock is not constant but varies with the shock angle. It

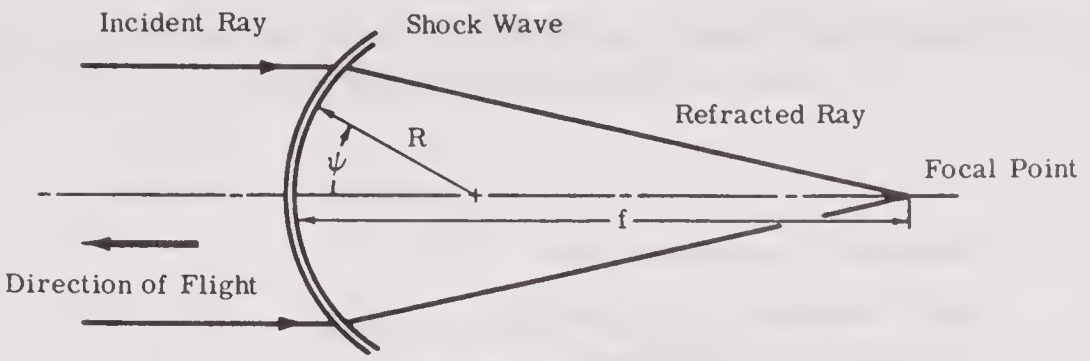


FIG. 21-20. Hemispherical shock wave refraction of incident radiation.

appears that each elemental annular ring, centered on the axis of symmetry, has its own focal length along the axis. Using the relations for the density change across the shock and Snell's law, this focal point can be computed as a function of the entrance angle, ψ. Figure 21-21 shows the results of such calculations. For Mach 3 flight at sea level, the focal length increases with ψ, but at Mach 5 it decreases with increasing ψ. Apparently, at some Mach number for sea-level flight the hemispherical shock behaves as a proper lens and has a single focal length. This Mach number appears to be about 3.26. The focal length becomes shorter as the Mach number increases because of the stronger shock. The simplified assumption of a spherical shock front obscures the detailed ray behavior due to the change of shock geometry with Mach number; however, this is of minor importance.

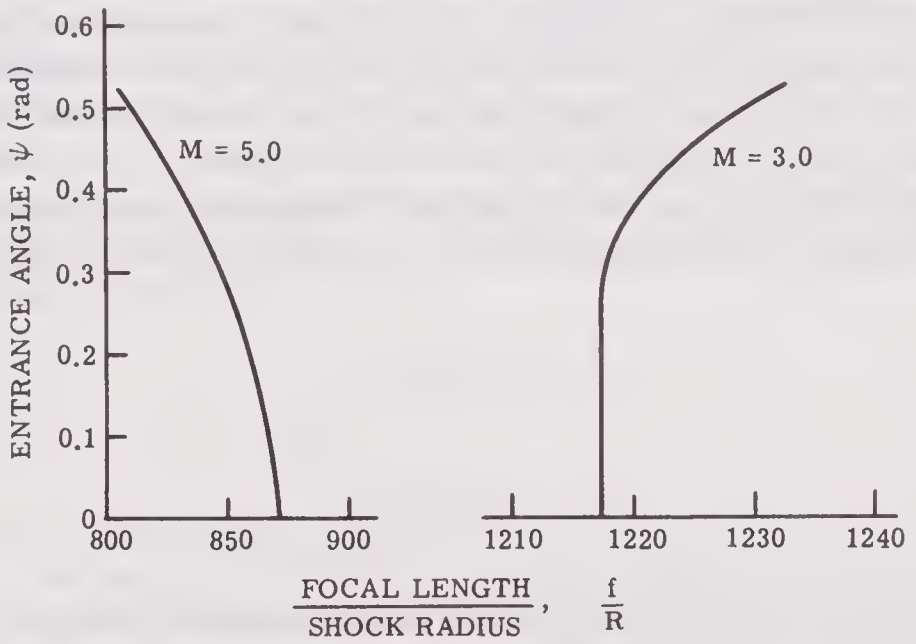


FIG. 21-21. Ratio of focal length to shock radius for hemispherical shock wave, $\lambda > 15\ \mu$, at sea level.

An additional optical element with a focal length equal to about a thousand times the shock radius is still significant. Consider, the same 15-cm, $f/1$ optical system behind a 25-cm-diameter hemispherical shock. Again, the geometrical growth of the point's image in the focal plane is compared to the theoretical limit of resolution of the system. Table 21-2 shows the effects of the hemispherical shock wave at Mach 3 and Mach 5. The geometrical blurring of the image by the shock in the uncorrected focal plane is seen to be 50 to 100 times greater than the circle of confusion. By moving the focal plane a few tenths of a millimeter to obtain an optimum focus in the presence of

TABLE 21-2. EFFECT OF HEMISPHERICAL SHOCK WAVE ON $f/1$, 15-CM OPTICAL SYSTEM

	$M = 3.0$ (*radians*)	$M = 5.0$ (*radians*)
Theoretical resolution limit	2.4×10^{-5}	2.4×10^{-5}
Angle subtended by image at 15-cm focus with shock	10^{-3}	1.4×10^{-3}
Angle subtended by image at optimum focus	1.5×10^{-5}	0.6×10^{-5}

the shock wave, the geometrical spreading of the image becomes less than the original circle of confusion. If the optical system looks off the axis of the shock system, the situation is far more complex. The density changes lose their axial symmetry, and the intercepted section of the shock wave is no longer hemispherical but more nearly parabolic.

21.3.3. Boundary-Layer Effects. The boundary layer can have any of several effects on the radiation passing through it. Steady-state effects can be similar to those found in the inviscid flow field, and the fluctuating nature of a turbulent boundary layer can have very serious optical consequences. Since the static pressure across a boundary layer is nearly constant, the variations in index of refraction are due entirely to the variation in static temperature. In the normal situation, n decreases from its free-stream value as the wall is approached. A minimum index of refraction would correspond to recovery temperature at the wall (Eq. 21-2), and this describes the "maximum effect" which can be expected (unless the wall is artificially heated by some means). It should be noted that boundary-layer thicknesses range from 0 in. at the leading edge of a sharp-edged surface to several inches well back along a full-scale aircraft. The approximate boundary-layer thickness δ on a flat plate is given in [2] as

$$\delta_{lam} = 5\sqrt{\nu x/u} \tag{21-28}$$

$$\delta_{turb} = 0.37x\left(\frac{ux}{\nu}\right)^{-1/5} \tag{21-29}$$

where u = velocity, ν = kinematic viscosity, and x = surface distance. It is seen that δ is proportional to $x^{1/2}$ in the laminar case and δ is proportional to $x^{4/5}$ in the turbulent case.

On the surface of a flat plate, two optical effects in the streamwise direction can be caused by this boundary-layer growth. First, the streamwise curvature of the boundary layer can cause a focusing effect, but some simple calculations show this to be negligible. At Mach 2 at sea level, for example, the focal length of the laminar boundary layer, only 30 cm from the leading edge, is greater than 10^{10} cm. The second effect is an angular deviation due to the locally nonparallel surfaces of the boundary layer (*i.e.*, the surface at the wall and the outer edge). Under the conditions mentioned above, the prismlike deviation is about 1 μrad. Increased velocity and altitude both decrease the magnitude of this effect.

The curvature of a body and the resulting curvature of the boundary layer can cause another type of focusing to take place. Considering only curvature and refraction in

a single plane, the apparent distances of an image from the circularly curved interface between two media of differing index of refraction can be written

$$\frac{n_1}{S_1} + \frac{n_2}{S_2} = \frac{n_2 - n_1}{R} \tag{21-30}$$

where R = radius of curvature of the interface
S_1 and S_2 = the apparent distances as seen from within the two media

If the boundary layer is assumed to be thin compared to the radius of curvature, (*i.e.,* the radius of curvature of the outer edge of the boundary is assumed equal to the inner radius), Eq. (21-30) can be used to determine the refractive effects of the boundary layer. When Eq. (21-30) is applied successively to thin layers of constant index of refraction, it is found that the indices of refraction at the wall and at the outer edge of the boundary layer will completely describe the refractive power of the boundary layer. By setting the original source distance at infinity, the incoming rays are made parallel, and the resulting apparent source distance from the boundary layer is the effective focal length. Thus,

$$\frac{f}{R} = \frac{n_w}{(n_w - n_\infty)} \tag{21-31}$$

The index of refraction in the local free stream, n_∞, can easily be determined once flight conditions have been established; however, the index of refraction at the wall, n_w, cannot be precisely predicted unless some knowledge of the wall temperature is at hand. As an example, the maximum effect will be calculated (*i.e.,* wall temperature equals recovery temperature). In general, the density of the boundary layer will be less at the wall than at the free-stream edge, and the minimum density possible is that associated with recovery temperature for a turbulent boundary layer. Recalling that the static pressure is constant across the boundary layer, Eq. (21-2) and the perfect gas law can be used to find the density difference. With these results and the Gladstone-Dale law, Eq. (21-31) can be rewritten

$$\frac{f}{R} = -\left(\frac{1}{k\rho_\infty}\right)(5r^{-1}M^{-2} + 1) - 5r^{-1}M^{-2} \tag{21-32}$$

From Eq. (21-32) it can be seen that the focal length is negative, which means that the boundary layer acts as a concave lens and makes incoming parallel rays diverge. In addition, the larger the absolute value of the ratio f/R, the less the effect that the boundary layer has on optical performance. At low Mach numbers the boundary layer is not a very important optical element, and at Mach numbers above 5 or 6 the effect approaches a limit of

$$\frac{f}{R} = -\frac{1}{k\rho_\infty} \tag{21-33}$$

Figure 21-22 shows the results of calculations based on Eq. (21-32) as a plot of f/R *versus* local free-stream Mach number for a cylindrical body traveling with its axis at zero angle of attack and incident parallel rays perpendicular to the axis. The strength of the boundary layer as an optical element can be seen to be about the same as that of the hemispherical and conical shock waves previously discussed.

The example shown in Fig. 21-22 is rather restricted. If the body under consideration is not cylindrical but is, for example, an ogive of circular cross section, the increasing

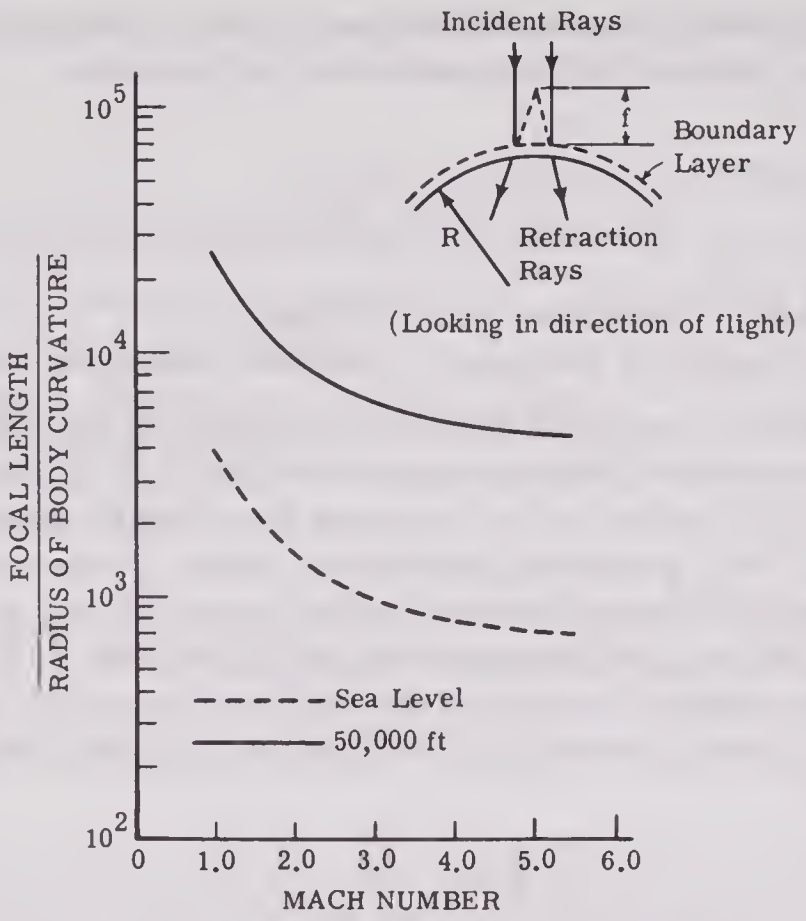


FIG. 21-22. Ratio of focal length to radius of boundary-layer curvature.

body radius causes the local flow properties to vary along the surface, and the variations in local free-stream Mach number and density may produce results somewhat different than those shown in Fig. 21-22. In addition, the curvature in two directions in this case would create considerable astigmatism.

It should be emphasized that the foregoing discussion describes the behavior of the radiation just before it passes from the boundary layer into the protective window or optical device. The geometry and properties of the outermost optical element must be considered when the effects of the boundary layer are calculated for a particular configuration. For instance, if a thin protective window with the same radius of curvature as the boundary layer separates the boundary layer from an evacuated cavity, the effective focal length of the boundary layer-window combination is

$$\frac{f}{R} = -\frac{1}{k\rho_\infty}$$

regardless of the flight speed or window material.

21.3.4. Fluctuating-Boundary-Layer Effects. The turbulent boundary layer is characterized by random variations in velocity superimposed on the mean velocity distribution. Integrated aerodynamic effects of the turbulent boundary layer, such as heat transfer and friction drag, are fairly well known; but little is known about the structure of the turbulence itself, and no entirely satisfactory mathematical model is available.

Several optical effects are present because of the nonuniformities in the boundary layer. The perturbations of the index of refraction causes the incoming rays to be refracted to a varying extent throughout the field; the local disturbances not being constant with time. When an image is focused by an optical system, these disturbances give rise to

(1) Image dancing—random shifting in apparent object position.

(2) Scintillation—fluctuation in image intensity.

(3) Loss in resolution—initially parallel rays not coming to focus at the same point in the image plane.

Several attempts to determine the magnitude of these effects have been carried out [16, 17, 18, and 19]. However, variations in experimental technique and measured quantities make it difficult to correlate the results. Some attempt to predict the loss in resolution for actual flight conditions was made with the following results [20]:

The scattering (loss in resolution) due to the turbulent boundary layer can be correlated on the basis of a parameter β', which is defined as [19]

$$\beta' = \frac{1}{\delta}\int_0^{\delta}\left[1-\frac{\rho(y)}{\rho_\infty}\right]dy \tag{21-34}$$

where δ = boundary-layer thickness
ρ_∞ = free-stream air density
$\rho(y)$ = local density as a function of distance from the wall (y)

To evaluate this integral, some model of the boundary-layer profile must be used. If one assumes that no heat is transferred to the wall, that the Prandtl number = 1, and that

$$\frac{u}{U_\infty} = \left(\frac{y}{\delta}\right)^{1/9}$$

the density ratio in air can be written

$$\frac{\rho}{\rho_\infty} = \left\{1 + 0.2M_\infty^2\left[1-\left(\frac{y}{\delta}\right)^{2/9}\right]\right\}^{-1} \tag{21-35}$$

In Fig. 21-23 the Stine and Winovich parameter $\beta'/\rho_\infty\delta$ is shown. The boundary-layer thickness on a flat plate, δ, is approximately

$$\delta = 0.37\,\frac{x}{Re^{1/5}} \tag{21-36}$$

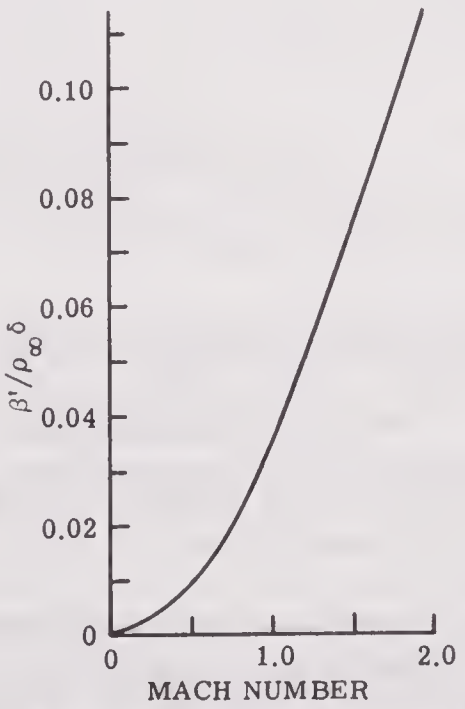


FIG. 21-23. Stine and Winovich scattering parameter *versus* Mach number for $u/U_\infty = (y/\delta)^{1/9}$.

From the data of [17, 18, 19] and the analyses of [18] and [20], it appears that for the wind tunnel experiments, the empirical relation is

$$\theta_B \approx KB\rho_\infty \sqrt{\delta} \tag{21-37}$$

where $K = 2.47$ rad ft$^{5/2}$/slug. The resolution limit, θ_B, is defined in [19] as the angular diameter of the aperture at the image plane through which 85% of the total energy from a point source (collimated beam) passes.

Equation (21-37) is very approximate and should be considered only as a guide to the general magnitude of the image degradation. The difficulties in making wind-tunnel measurements and the variations in technique among the various experimenters preclude an exact prediction of resolution loss. However, some valid general trends can be observed from Eq. (21-36). First, the resolution limit is proportional to the free-stream density (and therefore altitude). Second, the resolution loss increases approximately as $x^{2.5}$, x being the distance from the aerodynamic leading edge of the vehicle. Figure 21-24 shows Eq. (21-36) evaluated for several altitudes for a range of Mach numbers.

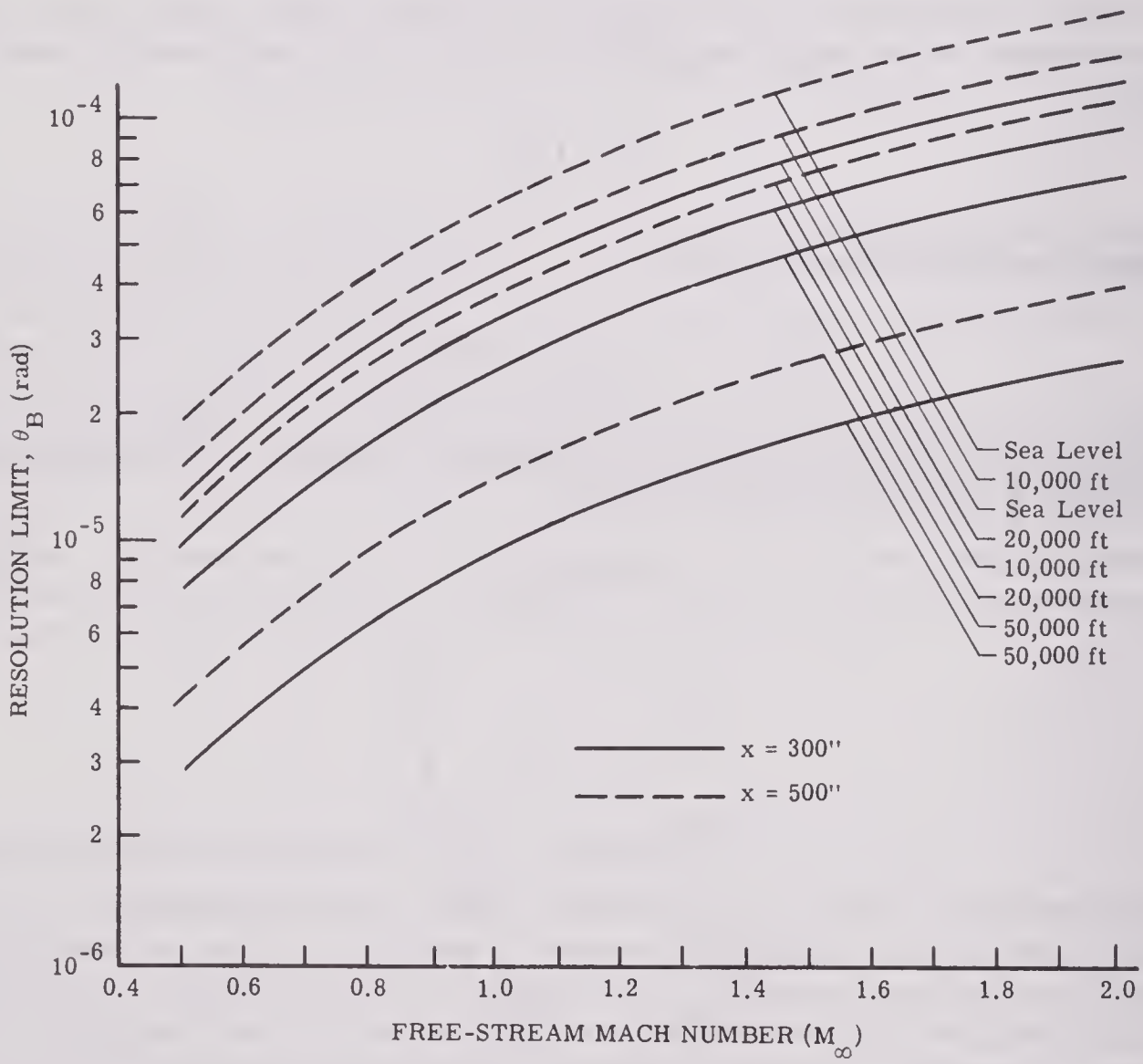


FIG. 21-24. Resolution limit as a function of M_∞ and body station.

In the treatment of this problem to date, no data have been obtained on the influence of aperture diameter. The larger-diameter astronomical telescopes are not as susceptible to atmospheric turbulence [21]. Tests so far have been made with apertures of 2 to 3 in.

21.4. Radiation from Heated Air

Most of the infrared radiation originating in the air surrounding the vehicle comes from heating of the air either by its passage through a strong shock wave or by its being accelerated in the boundary layer. Since the radiation energy associated with this flow field is many orders of magnitude below the kinetic and potential energies, the description of the flow field can be treated independently of the radiation. Once the flow-field properties are established, the radiation, in principle, can be estimated from laboratory measurements and from theory.

The calculation of the flow field is complex. Not only must the pressure and temperature be estimated at each point in the field from the vehicle out to the bow-shock wave, but also the nonequilibrium and chemical kinetic effects must be accounted for. These two considerations are vital if realistic estimates of air radiation are to be made, especially at high altitudes and high speeds. In particular the molecular vibration-rotation relaxation times may, or may not, be long enough to provide a radiation-free area in the forward part of the flow field. The constituents of high-temperature air (in particular CO and NO which do not occur at ambient conditions) must be found to calculate a reasonable spectrum.

If the fluid properties are known, the radiation characteristics can be estimated from standard works on gaseous radiation such as [22] and [23].

References

1. *High Speed Aerodynamics and Jet Propulsion, Vol. V. Turbulent Flows and Heat Transfer,* Princeton University Press, Princeton, N.J., 1959.
2. Schlichting, Hermann., *Boundary Layer Theory,* McGraw-Hill Book Co. Inc., New York, 1955.
3. Schneider, P. J., *Conduction Heat Transfer,* Addison-Wesley Publishing Co. Inc., Reading, Mass., 1955.
4. Fourier, J., *Analytical Theory of Heat,* Dover Publications Inc., New York, 1955.
5. Van Driest, E. R., *Investigation of Laminar Boundary Layer in Compressible Fluids Using the Crocco Method,* NACA Tech. Note 2597, Jan. 1952.
6. *Notes For a Special Summer Program in Aerodynamic Heating of Aircraft Structures in High Speed Flight,* Massachusetts Institute of Technology, Department of Aeronautical Engineering, 1956.
7. Sibulkin, M. J., *Aeronaut. Sci.,* Vol. 19, No. 570, 1952.
8. *Equations, Tables, and Charts for Compressible Flow,* NACA Tech. Rept. 1135 (1947).
9. Stine, H. A. and K. Wanlass, *Theoretical and Experimental Investigation of Aerodynamic Heating and Isothermal Heat Transfer Parameters on a Hemispherical Nose with Laminar Boundary Layer at Supersonic Mach Numbers,* NACA Tech. Note 3344, 1957.
10. Beckwith, I. E. and J. S. Gallagher, *Heat Transfer and Recovery Temperatures on a Sphere with Laminar Transitional and Turbulent Boundary Layers at Mach Numbers of 2.00 and 4.15,* NACA Tech. Note 4125, 1957.
11. Beardsley, N. F., "Infrared Transmitting Windows" (U), *Proc. of Infrared Information Symposium,* Vol. 1, No. 2, Dec. 1956 (CONFIDENTIAL).
12. Charwat, A. F., C. F. Dewey, J. N. Roos, and J. A. Hitz, "An Investigation of Separated Flows. Part II. Flow in the Cavity and Heat Transfer," *Journal of Aerospace Sciences,* Vol. 28, No. 7, July 1961.
13. Slater, J. C., and N. H. Frank, *Electromagnetism,* McGraw-Hill Book Co. Inc., New York, 1957.
14. *Handbook of Geophysics for Air Force Designers,* AFCRC, Cambridge, Mass., 1957.
15. Melkus, H. A., *The Behavior of a Light Ray Penetrating a Supersonic Flow Field,* U.S. Air Force, ARDC Report AFMDC-TR-59-39, Oct. 1959.
16. Liepmann, H. W., *Deflection and Diffusion of a Light Ray Passing Through a Boundary Layer,* Report SM-14397, Douglas Aircraft Company, Santa Monica, Calif., 1952.
17. Baskins, L. L. and L. E. Hamilton, *The Effect of Boundary Layer Thickness Upon the Optical Transmission Characteristics of a Supersonic Turbulent Boundary Layer,* Report NA1-54-756, Northrop Aircraft Co., Beverly Hills, Calif., Nov. 1954.
18. Bartlett, C.J., *The Scattering of Light Rays in a Supersonic Turbulent Boundary Layer.* M. S. Thesis, Massachusetts Institute of Technology, Department of Aeronautics and Astronautics, June 1961.

19. Stine, H.A., and W. Winovich, *Light Diffusion Through High-Speed Turbulent Boundary Layers,* NACA RM A56B21, May 1956.
20. Lorah, L. D., J. E. Nicholson, and R. E. Good, *Near-Field Aerodynamic Influences on Optical Systems* (U), Mithras Inc., Cambridge, Mass., Report MC-61-19-R-1. Sept. 1961 (CONFIDENTIAL).
21. Mikesell, A.H., *The Scintillation of Starlight,* Published by the U.S. Naval Observatory. Second Series, Vol. XVII, Part IV, 1955.
22. Penner, S. S., *Quantitative Molecular Spectroscopy and Gas Emissivities,* Addison-Wesley Publishing Co., Reading, Mass., 1959.
23. Plass, G. N., *Emissivity of the 4.3 Micron Band of Carbon Dioxide,* Scientific Report No. 2, Aeronutronic Systems Inc., Div. of Ford Motor Co. Newport Beach, Calif.

Chapter 22

PHYSICAL CONSTANTS AND CONVERSION FACTORS

J. A. Jenney and Richard Phillips

The University of Michigan

CONTENTS

22.1. Physical Constants 855
22.2. Length 855
22.3. Area 855
22.4. Volume 855
22.5. Angle 855
22.6. Mass 855
22.7. Density 856
22.8. Time 856
22.9. Velocity and Speed 856
22.10. Acceleration 856
22.11. Force 856
22.12. Torque 856
22.13. Pressure 856
22.14. Temperature 856
22.15. Work and Energy 857
22.16. Power 857
22.17. Electrical Units 857
22.18. Prefixes 857

Tables and Charts

Table 22-1. Physical Constants 858
Table 22-2. Length Conversions 860
Table 22-3. Equivalents of Common Fractions of Inches in Decimals and Millimeters 862
Fig. 22-1. Range Conversions 863
Table 22-4. Volume Conversions 864
Table 22-5. Angle Conversions 865
Fig. 22-2. Angle Conversion Chart 866
Table 22-6. Mass Conversions 866
Table 22-7. Useful Mass Units 867
Table 22-8. Density Conversions 867
Table 22-9. Time Conversions 868
Table 22-10. Velocity Conversions 869
Table 22-11. Acceleration Conversions 870
Table 22-12. Force Conversions 871
Table 22-13. Torque Conversions 871
Table 22-14. Pressure Conversions 872
Fig. 22-3. Temperature Scales 873
Fig. 22-4. High-Temperature Conversions 874
Table 22-15. Work and Energy Conversions 875
Fig. 22-5. Energy Conversions 876
Table 22-16. Spectroscopic Energy Conversions and Equivalences 878
Fig. 22-6. Spectroscopic Energy Conversions 878
Table 22-17. Power Conversions 879
Table 22-18. Electrical Unit Conversions 880
Table 22-19. Prefixes 881

22. Physical Constants and Conversion Factors

22.1. Physical Constants

Table 22-1(*A*) lists defined values and equivalents; Table 22-1(*B*) lists energy conversion factors; and Table 22-1(*C*) gives adjusted values of constants. The values are those recommended by the National Academy of Sciences and the National Research Council, as reported in *Physics Today,* issue of February 1964. The notation used follows that of this reference.

22.2. Length (l)

The standard of length is the meter, defined as 1,553,164.13 wavelengths of the cadmium red line in air at 760 mm pressure and 15°C. Conversions from one set of units to another are given in Table 22-2. Equivalents of some common fractions of inches in decimals and millimeters are given in Table 22-3. Figure 22-1 provides a graphical technique for length conversions in yards, kilometers, statute miles, and nautical miles.

22.3. Area (l^2)

Measurements of area are based on those of length; thus there is no standard area.

22.4. Volume (l^3)

Table 22-4 includes the common scientific units of volume.

22.5. Angle

An angle is usually defined in terms of a fraction of a circle, or the arc and radius of the arc. Given this way, it is dimensionless, although the name *radian* is usually specified. One radian is the angle subtended by an arc equal to the radius of the circle. When specified in degrees the angle is still dimensionless, but numerical difficulties arise in calculations because the number base 360 is now implicit. Conversions among degrees, radians, grades, etc., are given in Table 22-5. It should be noted that a mil is often defined in the same way as a milliradian, and is often given as 1/6400 of a circle, and also as a subtense (grade) of 1 part in 1000. These are all approximately equal for small angles. Here they are defined as follows:

1 milliradian (1 mrad) = 0.001 radian
1 angular mil (1 mil) = angle of 1 part in 1000
1 military mil (1 mmil) = 1/6400 circle

Figure 22-2 is a conversion chart for changing from milliradians to degrees and minutes for small angles.

22.6. Mass (m)

Mass is the quantity of matter. Table 22-6 provides data for converting from one mass unit to another. The following conversion for energy and mass equivalents is sometimes used:

$$1 \text{ g} = 5.61000 \times 10^{26} \text{ Mev}$$

The ratios of proton to electron mass is 1836.12, and the mass of the earth is 5.983×10^{24} kg (6.595×10^{21} tons).

Table 22-7 lists some useful mass units.

22.7. Density (ml^{-3})

Density is mass per unit volume. Specific gravity is the ratio of the density of a substance to that of water. In the metric system specific gravity and density have the same value; in the English system specific gravity must be multiplied by the density of water. Table 22-8 gives conversion factors. The density of water at 3.98°C and 760 mm pressure is 1 g/ml; the density of water at 4°C and 760 mm pressure is 1 g cm^{-3}. This accounts for the fact that 1 liter is equal to 1000.027 cc. The density of air at STP is 1.293×10^{-3} g cm^{-3}. Tables of the density of air as a function of pressure and temperature are available in standard handbooks.

22.8. Time (t)

The unit of time is defined as 1/86,400 of a mean solar day. Table 22-9 gives time conversion factors.

22.9. Velocity and Speed (lt^{-1})

Linear velocity is the time rate of motion in a fixed direction; angular velocity is the time rate of angular motion about an axis (t^{-1}). Table 22-10 presents factors for conversion among the different units of linear and angular velocity. Mach number is also a measure of speed and is defined as the ratio of the given speed to the speed of sound.

22.10. Acceleration (lt^{-2})

Acceleration is the time rate of change of velocity in speed or direction. Table 22-11 gives conversion factors for acceleration. The acceleration of gravity of the earth varies from 977.9 at the equator at sea level to 983.21 at the North Pole. Some other useful values are: Berlin, Germany, 981.26; London, England, 981.19; Madison, Wis., 980.35; New York, N.Y., 980.23; San Francisco, Calif., 979.94.

22.11. Force (mlt^{-2})

Force is usually defined in terms of mass and momentum. It is the quantity which imparts a change in momentum to a mass. Normally the famous equation $f = ma$ is used, but a more critical statement is $f = dp/dt$, where p is momentum. This allows for a change in mass [4,6]. Conversions are given in Table 22-12.

22.12. Torque (ml^2t^{-2})

The torque about an origin of a force acting at a point is the product of the length of the line between the point and the origin and the component of force perpendicular to the line. Table 22-13 presents torque conversion factors.

22.13. Pressure ($ml^{-1}t^{-2}$)

Pressure is force per unit area. One torr is the pressure of 1 mm Hg at 0°C and standard gravity. Table 22-14 provides pressure conversion factors.

22.14. Temperature

Temperature is a measure of the average translational kinetic energy of the molecules of a substance. The common temperature scales are Kelvin, Celsius (centigrade), Fahrenheit, and Rankine. The first three are given in Fig. 22-3 in a way that permits

direct conversion. For higher temperatures, Fig. 22-4 can be used for conversion from Kelvin to Rankine. The governing equations are:

$$°R = \frac{9}{5}°K$$

$$°C = \frac{5}{9}(°F - 32) = °K - 273$$

$$°F = \frac{9}{5}°C + 32 = °R - 459.69$$

22.15. Work and Energy (ml^2t^{-2})

Work is usually thought of as force acting through a distance. Energy is the capability of doing work. Both have the same dimensions. Conversions among units are given in Table 22-15. (See also Table 22-1*B*.) Energy conversions for the units most frequently used are given in Fig. 22-5. Spectroscopic energy conversions are given in Table 22-16, and a detailed conversion nomograph is given in Fig. 22-6. In the two nomographs, the columns implying equality of energy with frequency, with wave number, or with temperature are really equivalences. A dimensional proportionality factor h (Planck's constant) is implied in the one case and $1/2k$ (Boltzmann's constant) per degree of freedom in the other.

22.16. Power (ml^2t^{-3})

Power is the time rate at which work is done. Table 22-17 contains various units for power and conversion factors among them.

22.17. Electrical Units

Table 22-18 provides conversions among several of the common systems of units [4-6] used in measuring electromagnetic fields, currents, electric power, etc.

22.18. Prefixes

Some of the useful decade prefixes are given in Table 22-19.

References

1. Physics Today, **17** (Feb. 1964).
2. "Physical Constants and Conversion Factors," General Electric Company (1955).
3. "Range Conversion Chart," Aerojet-General Corporation (1958).
4. *Handbook of Chemistry and Physics,* Chemical Rubber Publishing Co., Cleveland, Ohio (1958).
5. D. C. Peasle, *Elements of Atomic Physics,* Prentice-Hall, New York (1955).
6. E. R. Cohen, K. N. Crowe, and J. W. M. Dumond, *Fundamental Constants of Physics,* Interscience, New York (1957).
7. W. A. Hiltner, *Astronomical Techniques,* U. of Chicago Press, Chicago (1960).

TABLE 22-1. PHYSICAL CONSTANTS

A. Defined Values and Equivalents

Meter	(m)	1 650 763.73 wavelengths in vacuo of the unperturbed transition $2p_{10} - 5d_5$ in ^{86}Kr
Kilogram	(kg)	mass of the international kilogram at Sèvres, France
Second	(s)	1/31 556 925.974 7 of the tropical year at 12^h ET, 0 January 1900
Degree Kelvin	(°K)	defined in the thermodynamic scale by assigning 273.16 °K to the triple point of water (freezing point, 273.15 °K = 0 °C)
Unified atomic mass unit	(u)	1/12 the mass of an atom of the ^{12}C nuclide
Mole	(mol)	amount of substance containing same number of atoms as 12g of pure ^{12}C
Standard acceleration of free fall	(g_n)	9.806 65 m s^{-2}, 980.665 cm s^{-2}
Normal atmospheric pressure	(atm)	101 325 N m^{-2}, 1 013 250 dyn cm^{-2}
Thermochemical calorie	(cal_{th})	4.1840 J, 4.1840×10^7 erg
International Steam Table calorie	(cal_{IT})	4.1868 J, 4.1868×10^7 erg
Liter	(l)	0.001 000 028 m^3, 1 000.028 cm^3 (recommended by CIPM, 1950)
Inch	(in.)	0.0254 m, 2.54 cm
Pound (avdp.)	(lb)	0.453 592 37 kg, 453.592 37 g

B. Energy Conversion Factors

	Formula	*Factor*	*Error limit*	*Conversion* — *Systeme International (MKSA)*	*Conversion* — *Centimeter-gram-second (CGS)*
Electron-volt	eV	1.60210	7	$\times10^{-19}$ J$(eV)^{-1}$	$\times10^{-12}$ erg $(eV)^{-1}$
Energy associated with					
Unified atomic mass unit	c^2/Ne	9.31478	15	10^8 $eV\ u^{-1}$	10^8 $eV\ u^{-1}$
Proton mass	m_pc^2/e	9.38256	15	10^8 $eV\ m_p^{-1}$	10^8 $eV\ m_p^{-1}$
Neutron mass	m_nc^2/e	9.39550	15	10^8 $eV\ m_n^{-1}$	10^8 $eV\ m_n^{-1}$
Electron mass	m_ec^2/e	5.11006	5	10^5 $eV\ m_e^{-1}$	10^5 $eV\ m_e^{-1}$
Cycle	e/h	2.41804	7	10^{14} $Hz(eV)^{-1}$	10^{14} $s^{-1}(eV)^{-1}$
Wavelength	ch/e	1.23981	4	10^{-6} eV m	10^{-4} eV cm
Wave number	e/ch	8.06573	23	10^5 $m^{-1}(eV)^{-1}$	10^3 $cm^{-1}(eV)^{-1}$
°K	e/k	1.16049	16	10^4 °K$(eV)^{-1}$	10^4 °K$(eV)^{-1}$

TABLE 22-1 (*Continued*). PHYSICAL CONSTANTS

C. Adjusted Values of Constants

Constant	*Symbol*	*Value*	*Est.‡ error limit*	*Unit* — *Système International (MKSA)*	*Unit* — *Centimeter-gram-second (CGS)*
Speed of light in vacuum	c	2.997925	3	$\times 10^{8}$ m s^{-1}	$\times 10^{10}$ cm s^{-1}
Elementary charge	e	1.60210	7	10^{-19} C	10^{-20} $cm^{1/2}g^{1/2}$*
		4.80298	20	.....	10^{-10} $cm^{3/2}g^{1/2}s^{-1}$†
Avogadro constant	N_A	6.02252	28	10^{23} mol^{-1}	10^{23} mol^{-1}
Electron rest mass	m_e	9.1091	4	10^{-31} kg	10^{-23} g
		5.48597	9	10^{-4} u	10^{-4} u
Proton rest mass	m_p	1.67252	8	10^{-27} kg	10^{-24} g
		1.00727663	24	10^{0} u	10^{0} u
Neutron rest mass	m_n	1.67482	8	10^{-27} kg	10^{-24} g
		1.0086654	13	10^{0} u	10^{0} u
Faraday constant	F	9.64870	16	10^{4} C mol^{-1}	10^{3} $cm^{1/2}g^{1/2}mol^{-1}$*
		2.89261	5	.....	10^{14} $cm^{3/2}g^{1/2}s^{-1}mol^{-1}$†
Planck constant	h	6.6256	5	10^{-34} J s	10^{-27} erg s
	h	1.05450	7	10^{-34} J s	10^{-27} erg s
Fine structure constant	α	7.29720	10	10^{-3}	10^{-3}
	$1/\alpha$	1.370388	19	10^{2}	10^{2}
	$\alpha/2\pi$	1.161385	16	10^{-3}	10^{-3}
	α^2	5.32492	14	10^{-5}	10^{-5}
Charge to mass ratio for electron	e/m_e	1.758796	19	10^{11} C kg^{-1}	10^{7} $cm^{1/2}g^{-1/2}$*
		5.27274	6	.....	10^{17} $cm^{3/2}g^{-1/2}s^{-1}$†
Quantum-charge ratio	h/e	4.13556	12	10^{-15} J s C^{-1}	10^{-7} $cm^{3/2}$ $g^{1/2}s^{-1}$*
		1.37947	4	.....	10^{-17} $cm^{1/2}$ $g^{1/2}$†
Compton wavelength of electron	λ_C	2.42621	6	10^{-12} m	10^{-10} cm
	$\lambda_C/2\pi$	3.86144	9	10^{-13} m	10^{-11} cm
Compton wavelength of proton	$\lambda_{C,p}$	1.32140	4	10^{-15} m	10^{-13} cm
	$\lambda_{C,\,p}/2$	2.10307	6	10^{-18} m	10^{-14} cm
Rydberg constant	R_∞	1.0973731	3	10^{7} m^{-1}	10^{5} cm^{-1}
Bohr radius	a_0	5.29167	7	10^{-11} m	10^{-9} cm
Electron radius	r_e	2.81777	11	10^{-15} m	10^{-13} cm
	r^2_e	7.9398	6	10^{-30} m^2	10^{-26} cm^2
Thomson cross section	$8\pi r^2_e/3$	6.6516	5	10^{-29} m^2	10^{-25} cm^2
Gyromagnetic ratio of proton	γ	2.67519	2	10^{8} rad $s^{-1}T^{-1}$	10^{4} rad $s^{-1}G^{-1}$*
	$\gamma/2\pi$	4.25770	3	10^{7} Hz T^{-1}	10^{3} $s^{-1}G^{-1}$*
(uncorrected for diamagnetism, H_2O)	γ'	2.67512	2	10^{8} rad $s^{-1}T^{-1}$	10^{4} rad $s^{-1}G^{-1}$*
	$\gamma'/2\pi$	4.25759	3	10^{7} Hz T^{-1}	10^{3} $s^{-1}G^{-1}$*
Bohr magneton	μ_B	9.2732	6	10^{-24} J T^{-1}	10^{-21} erg G^{-1}*
Nuclear magneton	μ_N	5.0505	4	10^{-27} J T^{-1}	10^{-24} erg G^{-1}*
Proton moment	μ_p	1.41049	13	10^{-26} J T^{-1}	10^{-23} erg G^{-1}*
	μ_p/μ_N	2.79276	7	10^{0}	10^{0} —
(uncorrected for diamagnetism, H_2O)	μ'_p/μ_N	2.79268	7	10^{0}	10^{0} —
Anomalous electron moment corrn	$(\mu_e/\mu_0) - 1$	1.159615	15	10^{-3}	10^{-3} —
Zeeman splitting constant	μ_B/hc	4.66858	4	10^{1} $m^{-1}T^{-1}$	10^{-5} $cm^{-1}G^{-1}$*
Gas constant	R	8.3143	12	10^{0} J $°K^{-1}$ mol^{-1}	10^{7} erg $°K^{-1}$ mol^{-1}
Normal volume perfect gas	V_0	2.24136	30	10^{-2} m^3 mol^{-1}	10^{4} cm^3 mol^{-1}
Boltzmann constant	k	1.38054	18	10^{-23} J $°K^{-1}$	10^{-16} erg $°K^{-1}$
First radiation constant ($2\pi hc^2$)	c_1	3.7405	3	10^{-18} W m^2	10^{-5} erg cm^2 s^{-1}
Second radiation constant	c_2	1.43879	19	10^{-2} m °K	10^{0} cm °K
Wien displacement constant	b	2.8978	4	10^{-3} m °K	10^{-1} cm °K
Stefan-Boltzmann constant	σ	5.6697	29	10^{-8} W m^{-2} $°K^{-4}$	10^{-5} erg cm^{-2} s^{-1} $°K^{-4}$
Gravitational constant	G	6.670	15	10^{-11} N m^2 kg^{-2}	10^{-8} dyn cm^2 g^{-2}

‡Based on 3 standard deviations, applied to last digits in preceding column.

*Electromagnetic system.

†Electrostatic system.

C – coulomb
J – joule
Hz – hertz
W – watt
N – newton
T – tesla
G – gauss

TABLE 22-2. LENGTH CONVERSIONS

		m	cm	km	μ	mμ	A	in.	mil
1 meter (m)	=	1	100	10^{-3}	10^{6}	10^{9}	10^{10}	39.3700	3.93700×10^{4}
1 centimeter (cm)	=	0.01	1	10^{-5}	10^{4}	10^{7}	10^{8}	0.393700	3.93700×10^{2}
1 kilometer (km)	=	1000	10^{5}	1	10^{9}	10^{12}	10^{13}	3.93700×10^{4}	3.93700×10^{7}
1 micron (μ)	=	10^{-6}	10^{-4}	10^{-9}	1	10^{3}	10^{4}	3.93700×10^{-5}	3.9370×10^{-2}
1 millimicron (mμ)	=	10^{-9}	10^{-7}	10^{-12}	10^{-3}	1	10	3.93700×10^{-8}	3.9370×10^{-5}
1 angstrom (Å or A)	=	10^{-10}	10^{-8}	10^{-13}	10^{-4}	0.1	1	3.93700×10^{-9}	3.93700×10^{-6}
1 inch (in.)	=	2.54×10^{-2}	2.540005	2.54×10^{-5}	2.54×10^{4}	2.54×10^{7}	2.54×10^{8}	1	10^{3}
1 mil	=	2.54×10^{-5}	2.54×10^{-3}	2.54×10^{-8}	25.400	2.54×10^{4}	2.54×10^{5}	10^{-3}	1
1 foot (ft)	=	0.304006	30.48006	3.048006×10^{-4}	3.048006×10^{5}	3.048006×10^{8}	3.048006×10^{9}	12	1.200×10^{4}
1 yard (yd)	=	0.91440	91.440	9.1440×10^{-4}	9.1440×10^{5}	9.1440×10^{8}	9.1440×10^{9}	36	3.600×10^{4}
1 fathom (fath)	=	1.828804	182.8804	1.828804×10^{-3}	1.828804×10^{6}	1.828804×10^{9}	1.828804×10^{10}	72	7.200×10^{4}
1 rod	=	5.0292	502.92	5.0292×10^{-3}	5.0292×10^{6}	5.0292×10^{9}	5.0292×10^{10}	198	1.98000×10^{5}
1 chain	=	20.1168	2011.68	2.01168×10^{-2}	2.01168×10^{7}	2.01168×10^{10}	2.01168×10^{11}	792	7.92000×10^{5}
1 link	=	0.201168	20.1168	2.01168×10^{-4}	2.01168×10^{5}	2.01168×10^{8}	2.01168×10^{9}	7.9200	7.92000×10^{3}
1 furlong (fur)	=	201.168	20116.8	0.201168	2.01168×10^{8}	2.01168×10^{11}	2.01168×10^{12}	7920.0	7.92000×10^{6}
1 (statute) mile (mi)	=	1.60935×10^{3}	1.60935×10^{5}	1.60935	1.60935×10^{9}	1.60935×10^{12}	1.60935×10^{13}	6.3360×10^{4}	6.3360×10^{6}
1 nautical mile (n mi) (International)	=	1852	1.8520×10^{5}	1.8520	1.8520×10^{9}	1.8520×10^{12}	1.8520×10^{13}	7.6338×10^{4}	7.6338×10^{6}
1 light year	=	9.4637×10^{15}	9.4637×10^{17}	9.4637×10^{12}	9.4637×10^{21}	9.4637×10^{24}	9.4637×10^{25}	3.9009×10^{17}	3.9009×10^{19}
1 parsec	=	3.0826×10^{16}	3.0826×10^{18}	3.0826×10^{13}	3.0826×10^{22}	3.0826×10^{25}	3.0826×10^{26}	1.2707×10^{18}	1.2707×10^{20}

TABLE 22-2 (*Continued*). LENGTH CONVERSIONS

ft	yd	fath	rod	chain	link	fur	mi	n mi	light yr	parsec
3.280833	1.093611	0.54681	0.19884	0.049710	4.9710	4.9710×10^{-3}	6.2137×10^{-4}	5.3996×10^{-4}	1.0556×10^{-18}	3.2438×10^{-17}
3.280833×10^{-2}	0.01093611	5.4681×10^{-2}	1.9884×10^{-2}	4.9710×10^{-4}	4.9710×10^{-2}	4.9710×10^{-5}	6.2137×10^{-6}	5.3996×10^{-6}	1.0566×10^{-18}	3.2438×10^{-19}
3280.833	1093.611	546.81	198.84	49.710	4971.0	4.9710	0.62137	0.53996	1.0566×10^{-13}	3.2438×10^{-14}
3.280833×10^{-6}	1.093611×10^{-6}	5.4681×10^{-7}	1.9884×10^{-7}	4.9710×10^{-8}	4.9710×10^{-6}	4.9710×10^{-9}	6.2137×10^{-10}	5.3996×10^{-10}	1.0566×10^{22}	3.2438×10^{-23}
3.280833×10^{-9}	1.093611×10^{-9}	5.4681×10^{-10}	1.9884×10^{-10}	4.9710×10^{-11}	4.9710×10^{-9}	4.9710×10^{-12}	6.2137×10^{-13}	5.3996×10^{-13}	1.0566×10^{-25}	3.2438×10^{-26}
3.280833×10^{-10}	1.093611×10^{-10}	5.4681×10^{-11}	1.9884×10^{-11}	4.9710×10^{-12}	4.9710×10^{-10}	4.9710×10^{-13}	6.2137×10^{-14}	5.3996×10^{-14}	1.0566×10^{-26}	3.2438×10^{-27}
0.08333	0.0277778	0.018889	5.0505×10^{-3}	1.2626×10^{-3}	1.2626×10^{-1}	1.2626×10^{-4}	1.5783×10^{-5}	1.3715×10^{-5}	2.6839×10^{-18}	8.2396×10^{-19}
8.3333×10^{-5}	2.77778×10^{-5}	1.8889×10^{-5}	5.0505×10^{-6}	1.2626×10^{-6}	1.2626×10^{-4}	1.2626×10^{-7}	1.5783×10^{-8}	1.3715×10^{-8}	2.6839×10^{-21}	8.2396×10^{-22}
1	0.3333	0.16667	6.0606×10^{-2}	0.01515	1.5152	1.5152×10^{-3}	1.89394×10^{-4}	1.6458×10^{-4}	3.2207×10^{-17}	9.8875×10^{-18}
3	1	0.5	0.18182	0.04545	4.5455	4.5455×10^{-3}	5.6819×10^{-4}	4.9375×10^{-4}	9.6622×10^{-17}	2.9663×10^{-17}
6	2	1	0.54545	0.09091	9.0909	9.0909×10^{-3}	1.1364×10^{-3}	9.8757×10^{-4}	1.9325×10^{-16}	5.9328×10^{-17}
16.500	5.5000	1.83333	1	0.25	25	0.025	3.125×10^{-3}	2.7156×10^{-3}	4.2574×10^{-16}	1.3070×10^{-16}
66	22	11	4	1	100	10^{-1}	1.2500×10^{-2}	1.0862×10^{-2}	2.1256×10^{-15}	6.5256×10^{-16}
0.66	0.22	0.11	0.04	10^{-2}	1	10^{-3}	1.2500×10^{-4}	1.0862×10^{-4}	2.1256×10^{-17}	6.5256×10^{-18}
660	220	110	40	10	10^{3}	1	0.12500	0.10862	2.1256×10^{-14}	6.5256×10^{-15}
5280	1760	880	320	80	8000	8	1	0.868979	1.7005×10^{-13}	5.2205×10^{-14}
6,076.10333	2025.37	1012.686	368.250	92.0624	9206.24	9.20624	1.15078	1	1.9569×10^{-13}	6.0077×10^{-14}
3.1049×10^{16}	1.0350×10^{16}	5.1750×10^{15}	1.8818×10^{15}	4.7405×10^{14}	4.7045×10^{16}	4.7045×10^{13}	5.8804×10^{12}	5.11000×10^{12}	1	0.307
1.0114×10^{17}	3.3713×10^{18}	1.6856×10^{16}	6.1296×10^{15}	1.5441×10^{15}	1.5441×10^{17}	1.5441×10^{14}	1.9154×10^{13}	1.6645×10^{13}	3.25733	1

TABLE 22-3. EQUIVALENTS OF COMMON FRACTIONS OF INCHES IN DECIMALS AND MILLIMETERS

Common Fractions	*Decimal*	*MM*	*Common Fractions*	*Decimal*	*MM*
1/64	0.0156	0.3969	33/64	.5156	13.0969
1/32	.0313	0.7938	17/32	0.5313	13.4938
3/64	.0469	1.1906	35/64	.5469	13.8906
1/16	.0625	1.5875	9/16	.5625	14.2875
5/64	.0781	1.9844	37/64	.5781	14.6844
3/32	.0938	2.3813	19/32	.5938	15.0813
7/64	.1094	2.7781	39/64	.6094	15.4781
1/8	.125	3.1750	5/8	.625	15.8750
9/64	.1406	3.5719	41/64	.6406	16.2719
5/32	.1563	3.9688	21/32	.6563	16.6688
11/64	.1719	4.3656	43/64	.6719	17.0656
3/16	.1875	4.7625	11/16	.6875	17.4625
13/64	.2031	5.1594	45/64	.7031	17.8594
7/32	.2188	5.5563	23/32	.7188	18.2563
15/64	.2344	5.9531	47/64	.7344	18.6531
1/4	.250	6.3500	3/4	.750	19.0500
17/64	.2656	6.7469	49/64	.7656	19.4469
9/32	.2813	7.1438	25/32	.7813	19.8438
19/64	.2969	7.5406	51/64	.7969	20.2406
5/16	.3125	7.9375	13/16	.8125	20.6375
21/64	.3281	8.3344	53/64	.8281	21.0344
11/32	.3438	8.7313	27/32	.8438	21.4313
23/64	.3594	9.1281	55/64	.8594	21.8281
3/8	.375	9.5250	7/8	.875	22.2250
25/64	.3906	9.9219	57/64	.8906	22.6219
13/32	.4063	10.3188	29/32	.9063	23.0188
27/64	.4219	10.7156	59/64	.9219	23.4156
7/16	.4375	11.1125	15/16	.9375	23.8125
29/64	.4531	11.5094	61/64	.9531	24.2094
15/32	.4688	11.9063	31/32	.9688	24.6063
31/64	.4844	12.3031	63/64	.9844	25.0031
1/2	.500	12.7000	1	1.000	25.4000

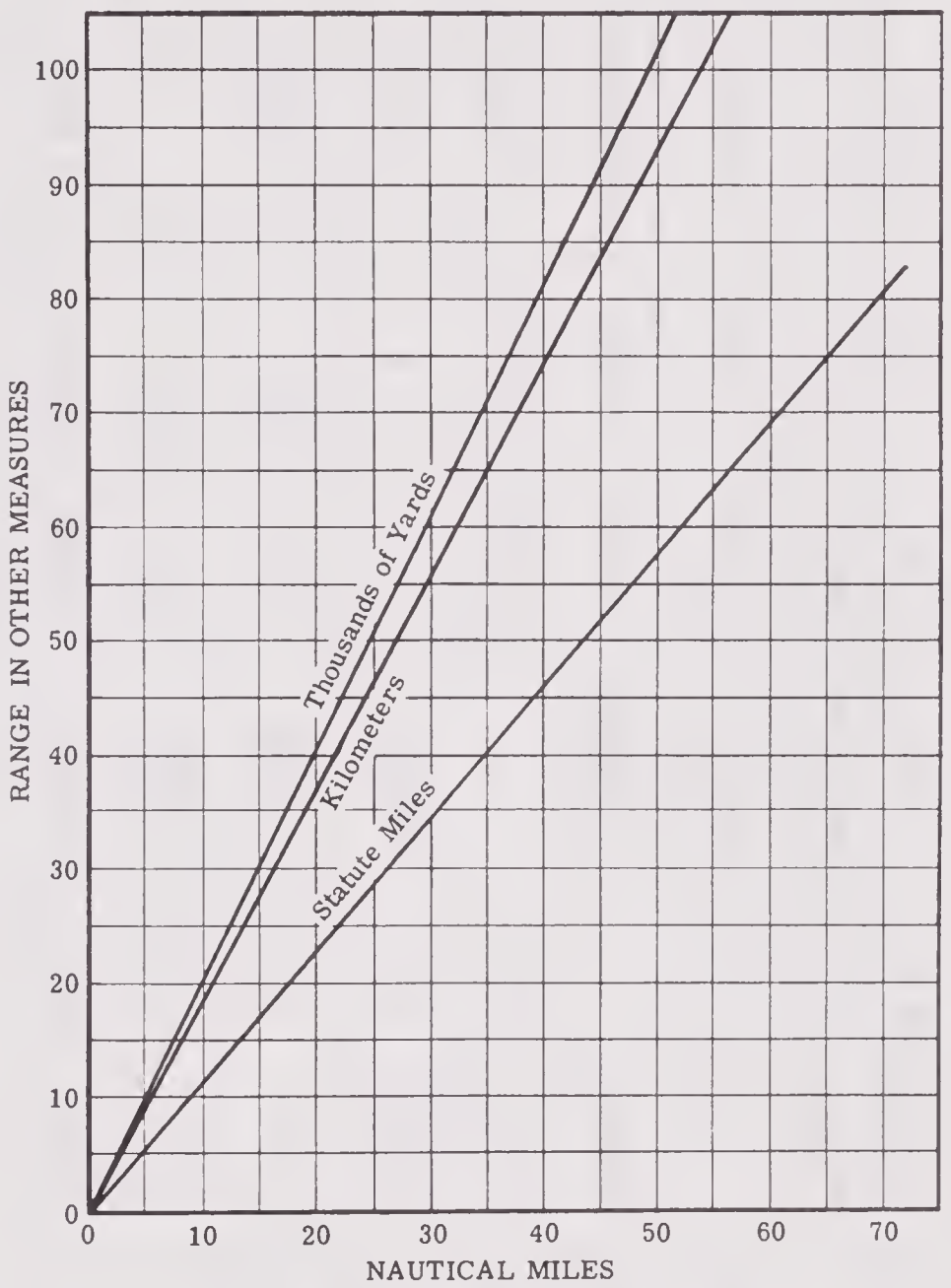


FIG. 22-1. Range Conversion Chart [3]

TABLE 22-4. VOLUME CONVERSIONS [2,4]

		cm^3	in^3	fl oz	liters	qt	pt	gal	ft^3
1 cubic centimeter (cm^3)	=	1	0.061023	0.03381	9.9997×10^{-4}	0.0010567	0.0021134	2.6417×10^{-4}	0.0003531
1 cubic inch (in^3)	=	16.387162	1	0.5541	0.0163868	0.017316	0.034632	0.0043290	0.0005787
1 fluid ounce (fl oz) (U.S.)	=	29.5737	1.80469	1	0.0295729	0.03125	0.0625	0.0078125	0.001044
1 liter	=	1000.027	61.025	33.8147	1	1.05671	2.1134	0.26418	0.3531
1 quart (qt) (liquid)	=	946.4	57.749	32	0.94633	1	2	0.25	0.0334
1 pint (pt) (liquid)	=	473.2	28.875	16	0.473167	0.5	1	0.125	0.0167
1 gallon (gal) (U.S.)	=	3785.4	231	128	3.78533	4	8	1	0.1336
1 cubic foot (ft^3)	=	28,317	1728	957.5	28.317	29.92	59.84	7.481	1

TABLE 22-5. ANGLE CONVERSIONS

		degree	min	sec	rad	mrad	circum-ference	quadrant	grade	centesimal min	centesimal sec
1 degree (°)	=	1	60	3600	1.7453×10^{-2}	1.7453×10^{-5}	2.7778×10^{-3}	1.11112×10^{-2}	1.11112×10^{1}	1.11112×10^{3}	1.11112×10^{5}
1 minute (′)	=	1.66667×10^{-2}	1	60	2.9089×10^{-4}	2.9089×10^{-7}	4.6296×10^{-5}	1.8518×10^{-4}	1.8518×10^{-2}	1.8518×10^{1}	1.8518×10^{3}
1 second (″)	=	2.77778×10^{-4}	1.66667×10^{-2}	1	4.8481×10^{-6}	4.8481×10^{-9}	7.71605×10^{-6}	3.0864×10^{-5}	3.0864×10^{-3}	3.0864×10^{-1}	3.0864×10^{1}
1 radian (rad)	=	5.72958×10^{1}	3.43775×10^{3}	2.06265×10^{5}	1	10^{3}	1.59155×10^{-1}	6.3662×10^{-1}	6.3662×10^{1}	6.3662×10^{3}	6.3662×10^{5}
1 milliradian (mrad)	=	5.72958×10^{-2}	3.43775	2.06265×10^{2}	10^{-3}	1	1.59155×10^{-4}	6.3662×10^{-4}	6.3662×10^{-2}	6.3662×10^{1}	6.3662×10^{3}
1 circumference	=	360	21,600	1,296,000	6.28319	6.28319×10^{3}	1	4	400	40,000	4.0000×10^{6}
1 quadrant	=	90	5,400	324,000	1.5708	1.5708×10^{3}	0.25000	1	10^{2}	10^{4}	10^{6}
1 grade	=	9.0000×10^{-1}	·54	3240	1.5708×10^{-2}	1.5708×10^{1}	2.5000×10^{-3}	10^{-2}	1	10^{2}	10^{4}
1 centesimal minute	=	9.0000×10^{-3}	5.40000×10^{-1}	3.24000×10^{1}	1.5708×10^{-4}	1.5708×10^{-1}	2.50000×10^{-5}	10^{-4}	10^{-2}	1	10^{2}
1 centesimal second	=	9.000×10^{-5}	5.4000×10^{-3}	3.24000×10^{-1}	1.5708×10^{-6}	1.5708×10^{-3}	2.5000×10^{-7}	10^{-6}	10^{-4}	10^{-2}	1

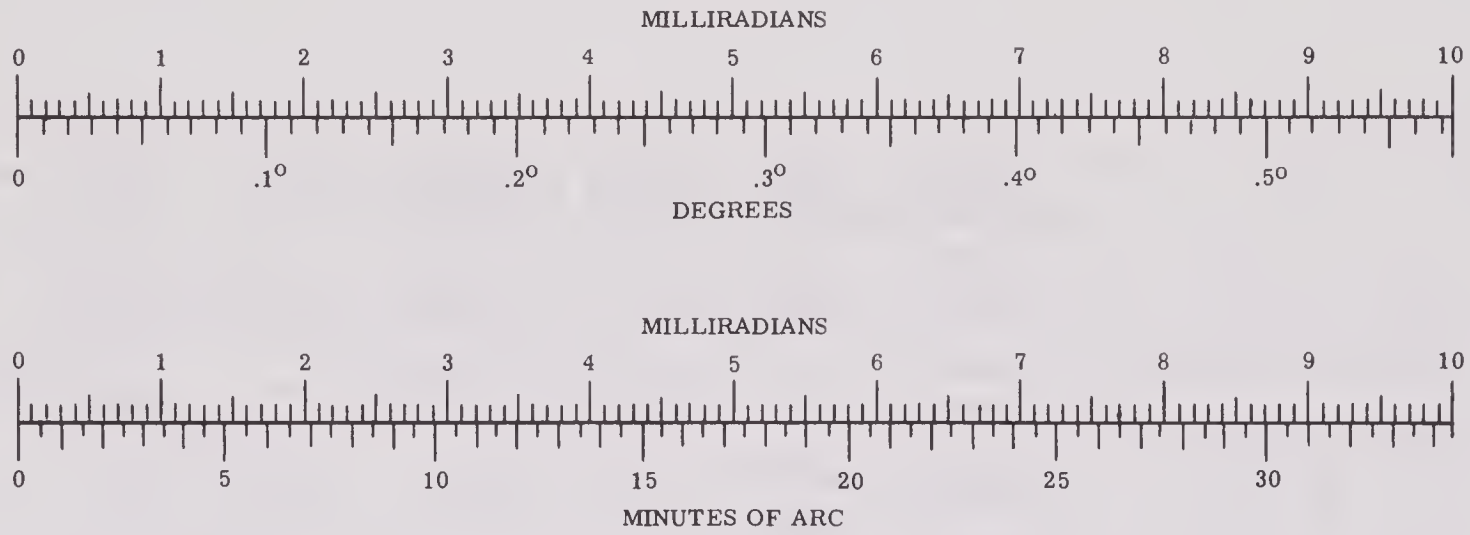


FIG. 22-2. Angle Conversion Chart

TABLE 22-6. MASS CONVERSIONS [4,6]

	mg	g	kg	ton	oz	lb	ton
1 milligram (mg) =	1	10^{-3}	10^{-6}	10^{-9}	35.2739×10^{-6}	2.20462×10^{-6}	1.10231×10^{-9}
1 gram (g) =	10^{3}	1	10^{-3}	10^{-6}	35.2739×10^{-3}	2.20462×10^{-3}	1.10231×10^{-6}
1 kilogram (kg) =	10^{6}	10^{3}	1	10^{-3}	35.2739	2.20462	1.10231×10^{-3}
1 ton (metric) =	10^{9}	10^{6}	10^{3}	1	35273.9	2204.62	1.10231
1 ounce (oz) =	28349.5	28.3495	0.0283495	28.3495×10^{-6}	1	0.0625	3.125×10^{-3}
1 pound (lb) =	453592	453.592	0.453592	0.453592×10^{-6}	16	1	0.0005
1 ton (avoirdupois) =	0.907185×10^{9}	907,185	907.185	0.907185	32,000	2000	1

TABLE 22-7. USEFUL MASS UNITS [4,6]

Quantity	*Valve* (*g*)
Mass of electron	9.1083×10^{-28}
Mass of proton	1.67239×10^{-24}
Mass of neutron	1.67470×10^{-24}
Atomic Mass Unit (Physical)	1.657×10^{-24}
Atomic Mass Unit (Chemical)	1.65746×10^{-24}

TABLE 22-8. DENSITY CONVERSIONS [4,6]

	g cm^{-3}	g $liter^{-1}$	lb in^{-3}	lb gal^{-1}	lb ft^{-3}
1 gram per cubic centimeter (g cm^{-3}) =	1	1000	0.03613	8.3452	62.43
1 gram per liter (g $liter^{-1}$) =	0.001	1	36.13×10^{-6}	8.3452×10^{-3}	0.06243
1 pound per cubic inch (lb in^{-3}) =	27.680	27680	1	231	1728
1 pound per gallon (lb gal^{-1}) =	0.11983	119.83	0.4329	1	7.481
1 pound per cubic foot (lb ft^{-3}) =	0.016018	16.018	5.787×10^{-4}	0.1337	1

TABLE 22-9. TIME CONVERSIONS (MEAN SOLAR TIME) [7]

	sec	msec	μsec	nsec	min	hr	day	yr
1 second (sec or ″) =	1	1000	10^{6}	10^{9}	1.6667×10^{-2}	2.7778×10^{-4}	1.1574×10^{-5}	3.1689×10^{-8}
1 millisecond (msec) =	10^{-3}	1	10^{3}	10^{6}	1.6667×10^{-5}	2.7778×10^{-7}	1.1574×10^{-8}	3.1689×10^{-11}
1 microsecond (μsec) =	10^{-6}	10^{-3}	1	10^{3}	1.6667×10^{-8}	2.7778×10^{-10}	1.1574×10^{-11}	3.1689×10^{-14}
1 nanosecond (nsec) =	10^{-9}	10^{-6}	10^{-3}	1	1.6667×10^{-11}	2.7778×10^{-13}	1.1574×10^{-14}	3.1689×10^{-17}
1 minute (min or ′) =	60	6.0000×10^{4}	6.0000×10^{7}	6.0000×10^{10}	1	1.6667×10^{-2}	6.9444×10^{-4}	1.9013×10^{-6}
1 hour (hr) =	3600	3.6000×10^{6}	3.6000×10^{9}	3.6000×10^{12}	60	1	4.1667×10^{-2}	1.1408×10^{-4}
1 day =	86400	8.64000×10^{7}	8.6400×10^{10}	8.6400×10^{13}	1440	24	1	2.7379×10^{-3}
1 year (yr) =	3.15569×10^{7}	3.15569×10^{10}	3.15569×10^{13}	3.15569×10^{16}	5.2596×10^{5}	8.7659×10^{3}	3.65242×10^{2}	1

Miscellaneous conversions:

1 sidereal year = 365.256 mean solar days = 8766.144 mean solar hours

1 calendar year = 365 mean solar days

1 leap year = 366 mean solar days

Table 22-10. Velocity Conversions

	cm sec^{-1}	m sec^{-1}	km sec^{-1}	in. sec^{-1}	ft sec^{-1}	mi sec^{-1}	n mi sec^{-1}	km hr^{-1}	in. hr^{-1}	ft hr^{-1}	mph	n mi hr^{-1}
1 centimeter per second (cm sec^{-1}) =	1	10^{-2}	10^{-5}	3.9370×10^{-1}	3.2808×10^{-2}	6.2137×10^{-6}	5.3996×10^{-6}	3.6000×10^{-2}	1.41732×10^{3}	1.1811×10^{2}	2.2369×10^{-2}	1.9438×10^{-2}
1 meter per second (m sec^{-1}) =	10^{2}	1	10^{-3}	3.9370×10^{1}	3.2808	6.2137×10^{-4}	5.3996×10^{-4}	3.6000	1.41732×10^{5}	1.1811×10^{4}	2.2369	1.9438
1 kilometer per second (km sec^{-1}) =	10^{5}	10^{3}	1	3.9310×10^{4}	3.2808×10^{3}	6.213×10^{-1}	5.3996×10^{-1}	3.6000×10^{3}	1.41732×10^{8}	1.1811×10^{7}	2.2369×10^{3}	1.9438×10^{3}
1 inch per second (in. sec^{-1}) =	2.54	2.54×10^{-2}	2.54×10^{-5}	1	8.3332×10^{-2}	1.5783×10^{-5}	1.3715×10^{-5}	9.1440×10^{-2}	3.6000×10^{3}	3.000×10^{2}	5.6817×10^{-2}	4.93725×10^{-2}
1 foot per second (ft sec^{-1}) =	3.0480×10^{1}	3.0480×10^{-1}	3.0480×10^{-4}	12	1	1.8940×10^{-4}	1.6458×10^{-4}	1.09728	4.3200×10^{4}	3.6000×10^{3}	6.81804×10^{-1}	5.9247×10^{-1}
1 (statute) mile per second (mi sec^{-1}) =	1.6093×10^{5}	1.6093×10^{3}	1.6093	6.3360×10^{4}	5280	1	8.6898×10^{-1}	5.7936×10^{3}	2.28096×10^{8}	1.9008×10^{7}	3.6000×10^{3}	3.12824×10^{3}
1 nautical mile per second (n mi sec^{-1}) =	1.85195×10^{5}	1.85195×10^{3}	1.85195	7.29134×10^{4}	6,076.1033	1.15078	1	6.67187×10^{3}	2.62488×10^{8}	2.1874×10^{7}	4.1428×10^{3}	3.6000×10^{3}
1 kilometer per hour (km hr^{-1}) =	2.7758×10^{1}	2.7758×10^{-1}	2.7758×10^{-4}	1.09287×10^{1}	9.1072×10^{-1}	1.72486×10^{-4}	1.49886×10^{-4}	1	3.9370×10^{4}	3.2786×10^{3}	6.20948×10^{-1}	5.3959×10^{-1}
1 inch per hour (in. hr^{-1}) =	7.05053×10^{-4}	7.05053×10^{-6}	7.05053×10^{-9}	2.7759×10^{-4}	2.3132×10^{-5}	4.38114×10^{-9}	3.8071×10^{-9}	2.54000×10^{-5}	1	8.3333×10^{-2}	1.57720×10^{-5}	1.37056×10^{-5}
1 foot per hour (ft hr^{-1}) =	8.46064×10^{-3}	8.46064×10^{-5}	8.46064×10^{-8}	3.3311×10^{-5}	2.77584×10^{-4}	5.25737×10^{-8}	4.56852×10^{-8}	3.0480×10^{-4}	12	1	1.89394×10^{-4}	1.64467×10^{-4}
1 (statute) mile per hour (mph) =	4.4672×10^{1}	4.4672×10^{-1}	4.4672×10^{-4}	1.7588×10^{-1}	1.4667	2.7759×10^{-4}	2.4122×10^{-4}	1.6093	6.3360×10^{4}	5280	1	8.6889×10^{-1}
1 nautical mile per hour (n mi hr^{-1}) =	5.1407×10^{1}	5.1407×10^{-1}	5.1407×10^{-4}	2.02397×10^{-1}	1.6878	3.1944×10^{-4}	2.7759×10^{-4}	1.8519	7.2913×10^{4}	6.07606×10^{3}	1.15078	1

Table 22-11. Acceleration Conversions

	m sec^{-2}	m min^{-2}	m hr^{-2}	m hr^{-1} sec^{-1}	in. sec^{-2}	in. min^{-2}	in. hr^{-2}	ft sec^{-2}	ft min^{-2}	ft min^{-1} sec^{-1}	mi sec^{-2}	mi min^{-2}	mi hr^{-2}	mi hr^{-1} sec^{-1}
1 meter per second per second (m sec^{-2}) =	1	3.6000 × 10^{3}	1.29600 × 10^{7}	3.6000 × 10^{3}	3.9370 × 10^{1}	1.41732 × 10^{5}	5.10235 × 10^{8}	3.28083	1.18110 × 10^{4}	1.9685 × 10^{2}	6.21373 × 10^{-4}	2.23694	8.05299 × 10^{3}	2.23694
1 meter per minute per minute (m min^{-2}) =	2.77778 × 10^{-4}	1	3.6000 × 10^{3}	1	1.09361 × 10^{-2}	3.93700 × 10^{1}	1.41732 × 10^{5}	9.11343 × 10^{-4}	3.28083	5.46806 × 10^{-2}	1.72604 × 10^{-7}	6.21373 × 10^{-4}	2.23694	6.21373 × 10^{-4}
1 meter per hour per hour (m hr^{-2}) =	7.71605 × 10^{-8}	2.77778 × 10^{-4}	1	2.77778 × 10^{-4}	3.03781 × 10^{-6}	1.09361 × 10^{-2}	3.93700 × 10^{1}	2.53151 × 10^{1}	9.11343 × 10^{-4}	1.51890 × 10^{-5}	4.79454 × 10^{-11}	1.72604 × 10^{-7}	6.21373 × 10^{-4}	1.72604 × 10^{-7}
1 meter per hour per second (m hr^{-1} sec^{-1}) =	2.77778 × 10^{-4}	1	3.60000 × 10^{3}	1	1.09361 × 10^{-2}	3.9370 × 10^{1}	1.41732 × 10^{5}	9.11343 × 10^{-4}	3.28083	5.46806 × 10^{-2}	1.726035 × 10^{-7}	6.21373 × 10^{-4}	2.23694	6.21373 × 10^{-4}
1 inch per second per second (in. sec^{-2}) =	2.54000 × 10^{-2}	9.14400 × 10^{1}	3.29184 × 10^{5}	9.14400 × 10^{1}	1	3.6000 × 10^{3}	1.29600 × 10^{7}	8.33333 × 10^{-2}	3.00000 × 10^{2}	5.00000	1.57829 × 10^{-5}	5.68183 × 10^{-2}	2.04546 × 10^{2}	5.60966 × 10^{-2}
1 inch per minute per minute (in. min^{-2}) =	7.055556 × 10^{-6}	2.54000 × 10^{-2}	9.14400 × 10^{1}	2.5400 × 10^{-2}	2.77778 × 10^{-4}	1	3.60000 × 10^{3}	2.31481 × 10^{-5}	8.33333 × 10^{-2}	1.38889 × 10^{-3}	4.38413 × 10^{-9}	1.57829 × 10^{-5}	5.68183 × 10^{-2}	1.57829 × 10^{-5}
1 inch per hour per hour (in. hr^{-2}) =	1.95988 × 10^{-9}	7.05556 × 10^{-6}	2.54000 × 10^{-2}	7.05556 × 10^{-6}	7.71605 × 10^{-8}	2.77778 × 10^{-4}	1	6.43004 × 10^{-9}	2.31481 × 10^{-5}	3.85803 × 10^{-7}	1.21781 × 10^{-12}	4.38413 × 10^{-9}	1.57829 × 10^{-5}	4.38413 × 10^{-9}
1 foot per second per second (ft sec^{-2}) =	3.04800 × 10^{-1}	1.09728 × 10^{3}	3.950208 × 10^{6}	1.09728 × 10^{3}	12	4.32000 × 10^{4}	1.55520 × 10^{8}	1	3.60000 × 10^{3}	6.00000 × 10^{1}	1.89394 × 10^{-4}	6.81820 × 10^{-1}	2.45455 × 10^{3}	6.81820 × 10^{-1}
1 foot per minute per minute (ft min^{-2}) =	8.46667 × 10^{-5}	3.04800 × 10^{-1}	1.09728 × 10^{3}	3.04800 × 10^{-1}	3.33333 × 10^{-3}	12	4.32000 × 10^{4}	2.77778 × 10^{-4}	1	1.66667 × 10^{-2}	5.26096 × 10^{-8}	1.89394 × 10^{-4}	6.81820 × 10^{-1}	1.89394 × 10^{-1}
1 foot per minute per second (ft min^{-1} sec^{-1}) =	5.08000 × 10^{-3}	1.82880 × 10^{1}	6.58368 × 10^{4}	1.82880 × 10^{1}	2.0000 × 10^{-1}	7.20000 × 10^{2}	2.5920 × 10^{6}	1.66667 × 10^{-2}	60	1	3.15657 × 10^{-6}	1.13637 × 10^{-2}	4.09092 × 10^{1}	1.13637 × 10^{-2}
1 mile per second per second (mi sec^{-2}) =	1.60934 × 10^{3}	5.79364 × 10^{6}	2.71931 × 10^{10}	5.79364 × 10^{6}	6.33600 × 10^{4}	2.28096 × 10^{8}	8.211456 × 10^{11}	5.28000 × 10^{3}	1.90080 × 10^{7}	3.16800 × 10^{5}	1	3.6000 × 10^{3}	1.29600 × 10^{7}	3.6000 × 10^{3}
1 mile per minute per minute (mi min^{-2}) =	4.47040 × 10^{-1}	1.60934 × 10^{3}	7.55364 × 10^{6}	1.60934 × 10^{3}	1.76000 × 10^{1}	6.33600 × 10^{4}	2.28096 × 10^{8}	1.46667	5.28000 × 10^{3}	8.8000 × 10^{1}	2.77778 × 10^{-4}	1	3.6000 × 10^{3}	1
1 mile per hour per hour (mi hr^{-2}) =	1.241778 × 10^{-4}	4.47040 × 10^{-1}	2.09823 × 10^{3}	4.47040 × 10^{-1}	4.88889 × 10^{-3}	1.76000 × 10^{1}	6.33600 × 10^{4}	4.07407 × 10^{-4}	1.46667	2.44444 × 10^{-2}	7.71605 × 10^{-8}	2.77778 × 10^{-4}	1	2.77778 × 10^{-4}
1 mile per hour per second (mi hr^{-1} sec^{-1}) =	4.47040 × 10^{-1}	1.60934 × 10^{3}	7.55364 × 10^{6}	1.60934 × 10^{3}	1.76000 × 10^{1}	6.33600 × 10^{4}	2.28096 × 10^{8}	1.46667	5.28000 × 10^{3}	8.80000 × 10^{1}	2.77778 × 10^{-4}	1	3.60000 × 10^{3}	1

TABLE 22-12. FORCE CONVERSION* [4,6]

		dyne	newton	g wt	kg wt	poundal	lb wt	ton wt
1 dyne	=	1	10^5	0.0010197	1.0197×10^{-6}	7.2330×10^{-5}	2.2481×10^{-6}	1.1241×10^{-9}
1 newton	=	10^5	1	101.97	0.10197	7.2330	22.481	1.1241×10^{-4}
1 gram weight (g wt)	=	980.665	0.0980665	1	0.001	0.070932	2.2046×10^{-3}	0.11023×10^{-5}
1 kilogram weight (kg wt)	=	980665	9.80665	1000	1	70.9032	2.2046	0.11023×10^{-2}
1 poundal	=	1.3825×10^4	0.13825	14.098	0.014098	1	0.031081	1.5492×10^{-5}
1 pound weight (lb wt)	=	4.4482×10^5	4.4482	453.59	0.45359	32.174	1	0.0005
1 ton weight (2000 lb)	=	8.8964×10^8	8896.4	907180	907.180	64548	2000	1

*Conversion between gravitational and absolute units of force are based on the values of g of 980.665 cm sec^{-2} and 32.174 ft sec^{-2}.

TABLE 22-13. TORQUE CONVERSIONS

		dyne cm	kg m	lb wt ft	lb wt in.	poundal ft
1 dyne centimeter (dyne cm)	=	1	1.0197×10^{-8}	7.3757×10^{-8}	8.8511×10^{-7}	2.3731×10^{-6}
1 kilogram meter (kg-m)	=	9.8066×10^7	1	7.23305	86.799	232.72
1 pound weight foot (lb wt ft)	=	1.3558×10^7	0.1383	1	12	32.174
1 pound weight inch (lb wt in.)	=	1.1298×10^6	0.01152	0.08333	1	2.6811
1 poundal foot (poundal ft)	=	4.2140×10^5	4.2970×10^{-3}	0.03108	0.37299	1

TABLE 22-14. PRESSURE CONVERSIONS [4,6]

		dyne cm^{-2}	g cm^{-2}	kg m^{-2}	mm Hg	atm	lb $in.^{-2}$	lb ft^{-2}	ton $in.^{-2}$	ton ft^{-2}
1 dyne per square centimeter (dyne cm^{-2})	=	1	0.0010197	0.010197	7.5006×10^{-4}	9.8692×10^{-7}	1.4504×10^{-5}	0.0020886	0.7252×10^{-8}	1.0443×10^{-6}
1 gram per square centimeter (g cm^{-2})	=	980.665	1	10	0.73556	9.6784×10^{-4}	0.014223	2.0482	7.112×10^{-6}	1.0241×10^{-3}
1 kilogram per square meter (kg m^{-2})	=	98.0665	0.1	1	0.073556	9.6784×10^{-5}	0.0014223	0.20482	7.112×10^{-7}	1.0241×10^{-4}
1 millimeter of mercury (mm Hg)	=	1,333.22	1.3595	13.595	1	0.001315	0.019337	2.7845	8.668×10^{-6}	1.3923×10^{-3}
1 atmosphere (atm)	=	1.01325×10^{6}	1033.2	1.0332×10^{4}	760	1	14.696	21,162	0.0073480	1.0581
1 pound per square inch (lb $in.^{-2}$)	=	68,947	70.307	703.07	51.715	0.068046	1	144	5×10^{-4}	0.0720
1 pound per square foot (lb ft^{-2})	=	478.80	0.48824	4.8824	0.35713	4.7254	0.0064444	1	3.4723×10^{-6}	5×10^{-4}
1 ton per square inch (ton $in.^{-2}$)	=	1.3789×10^{8}	140,620	1.406200	103,430	136.092	2,000	288,000	1	144
1 ton per square foot (ton ft^{-2})	=	9.5760×10^{5}	976.48	9,764.8	71,826	0.94509	13.889	2,000	0.0064444	1

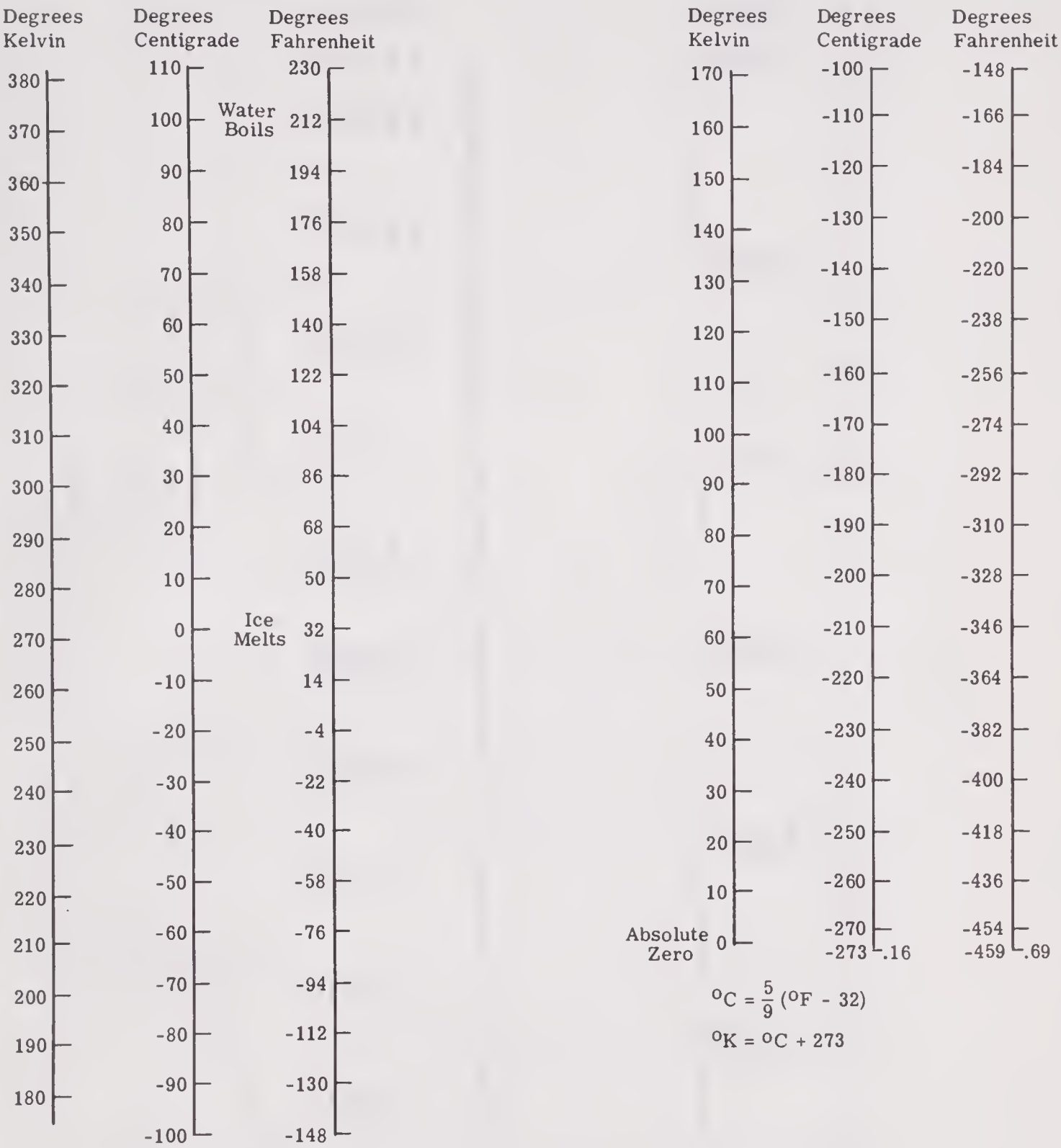


FIG. 22-3. Temperature Scales

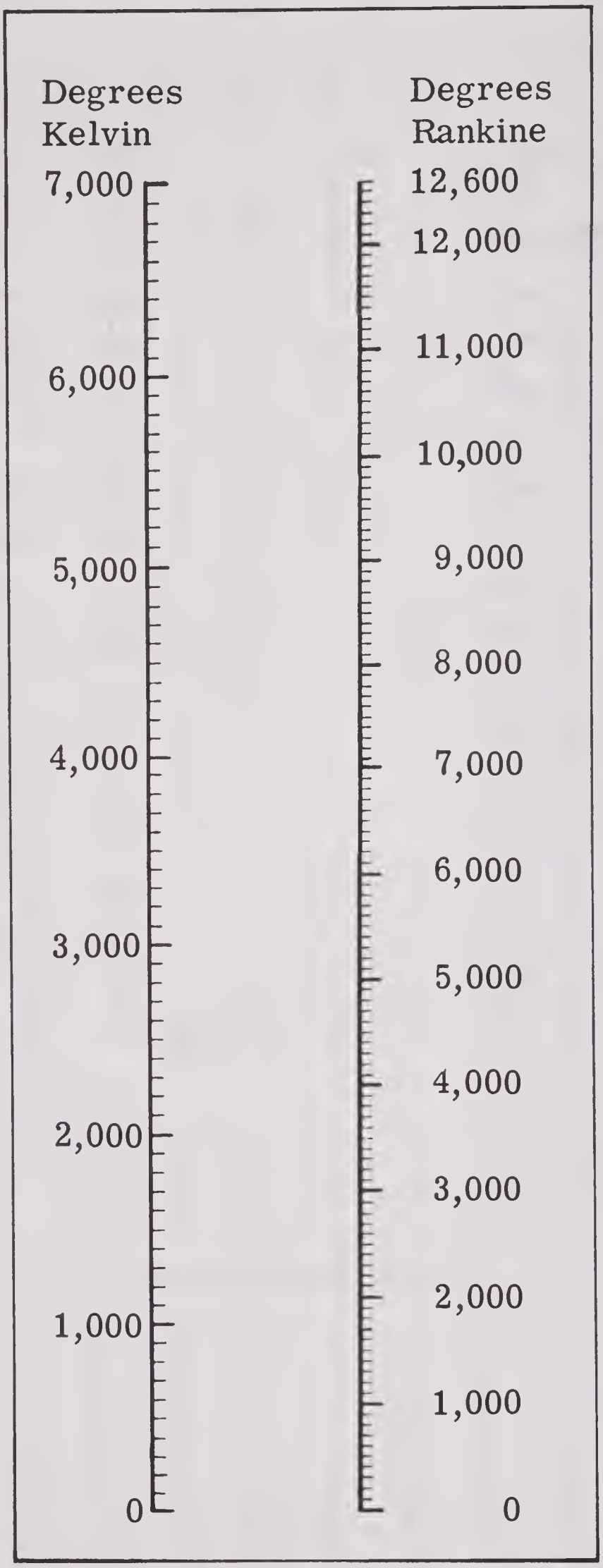


FIG. 22-4. High-Temperature Conversions [3]

Table 22-15. Work and Energy Conversions [2,4]

		joule (absolute)	joule (international)	erg	Btu (mean)	kwh	g cal (mean)	ev	ft lb	ft poundal	g cal (mean)	hp hr	g cal	liter atm	electronic charge	kg m
1 joule (absolute)	=	1	0.999835	10^{7}	9.4805×10^{-4}	2.778×10^{-7}	0.23889	0.62419×10^{19}	0.73756	23.730	2.3889×10^{-4}	3.725×10^{-7}	0.238948	9.869×10^{-3}	0.62786×10^{19}	0.10197
1 joule (international)	=	1.000165	1	1.000165×10^{7}	9.4821×10^{-4}	2.778×10^{-7}	0.23893	0.62429×10^{19}	0.73768	41.44218	2.38×10^{-4}	0.372417×10^{-6}	0.238959	9.87×10^{-3}	6.27968×10^{18}	0.101988
1 erg	=	1×10^{-7}	0.999853 $\times 10^{7}$	1	9.4805×10^{-11}	2.778×10^{-14}	2.388×10^{-8}	7.3756×10^{-9}	7.3736×10^{-8}	2.37303×10^{-6}	1.4333×10^{-9}	1.3412×10^{-10}	2.3892×10^{-8}	9.868×10^{-10}	6.27864×10^{-13}	1.01971×10^{-8}
1 British thermal unit (mean) (Btu)	=	1054.8	1054.625958	1.0548×10^{10}	1	2.930×10^{-4}	251.98	6.5843×10^{21}	777.97	2.5030×10^{4}	0.25198	3.9292×10^{-4}	453.720508	10.409	6.623×10^{21}	107.56
1 kilowatt hour (kwh)	=	3.6000×10^{6}	3.599406×10^{6}	3.6000×10^{13}	3413.0	1	8.6001×10^{5}	2.24691×10^{25}	2.6552×10^{6}	8.542335×10^{7}	860.585197	1.3410	8.6004×10^{5}	3.552×10^{4}	2.26014×10^{25}	3.6710×10^{5}
1 gram calorie (mean) (g cal)	=	4.186	4.185313	4.186×10^{7}	0.0039685	1.163×10^{-6}	1	2.612944×10^{19}	3.0874	99.334	1×10^{-3}	1.5593×10^{-6}	1.000236	0.041311	2.62822×10^{19}	0.42685
1 electron volt (ev)	=	1.60207×10^{-19}	1.601805×10^{-19}	1.60207×10^{-12}	1.5188×10^{-22}	4.451×10^{-26}	3.8272×10^{-20}	1	1.1816×10^{-19}	38.01712×10^{-19}	3.82718×10^{-23}	5.9677×10^{-26}	0.382811×10^{-19}	15.8108×10^{-22}	1.00587	0.16336×10^{-16}
1 foot pound (ft lb)	=	1.35582	1.355596	1.35582×10^{7}	0.0012854	3.7662×10^{-7}	0.32389	8.4628×10^{18}	1	32.174	3.2389×10^{-4}	5.0505×10^{-7}	0.32397	0.013381	0.85126×10^{19}	0.138255
1 foot poundal (ft poundal)	=	0.04	0.042133	4.21402×10^{5}	3.9952×10^{-5}	1.171×10^{-7}	0.010067	2.6302×10^{17}	0.031081	1	0.100668×10^{-4}	0.15697×10^{-7}	0.010069	4.1589×10^{-4}	2.6458×10^{21}	0.0042972
1 kilogram calorie (mean) (kg cal)	=	4186	4185.28838	4.186×10^{10}	3.9685	0.0011628	1000	$2612.85593 \times 10^{19}$	3087.4	99333.78	1	0.0015593	1000	41.3116	2.6288×10^{16}	426.85
1 horsepower hour (hp hr)	=	2.6845×10^{6}	2.685156×10^{6}	2.6845×10^{13}	2545.0	0.7457	6.4130×10^{5}	1.67324×10^{25}	1.9800×10^{6}	63.72928×10^{6}	6.41562×10^{2}	1	0.64171	26.5041×10^{3}	1.68618×10^{25}	2.7374×10^{5}
1 gram calorie (15°C) (g cal)	=	4.1855	4.184809	4.1855×10^{7}	0.003968	1.163×10^{-6}	0.99987	2.61254×10^{19}	3.0871	99.3219	9.9987×10^{-4}	15.5909×10^{-7}	1	41.3066×10^{-3}	2.6279×10^{19}	2.7385×10^{7}
1 liter atmosphere (normal) (liter atm)	=	101.328	101.311280	101.328×10^{7}	0.09607	2.815×10^{-5}	24.206	63.24792×10^{19}	74.735	2404.5	242.0624×10^{-4}	3.7745×10^{-5}	24.21212	1	63.61979×10^{19}	10.333
1 International volt (electronic charge)	=	1.5927×10^{-19}	1.592437×10^{-19}	1.5927×10^{-12}	1.5100×10^{-22}	4.425×10^{-26}	3.8048×10^{-20}	0.99414	1.1747×10^{-19}	37.7994×10^{-19}	3.8048×10^{-23}	5.9328×10^{-26}	0.38057×10^{-19}	15.7183×10^{-22}	1	1.624×10^{-20}
1 kilogram meter (kg m)	=	9.80665	9.805031	9.80665×10^{7}	0.0092972	2.7235×10^{-6}	2.3427	6.12121×10^{19}	7.2330	232.71	23.4271×10^{-4}	3.6530×10^{-6}	2.3427	0.096782	6.1572×10^{19}	1

1 joule (absolute) = 1 newton m = 1 w sec
1 erg = 1 dyne cm

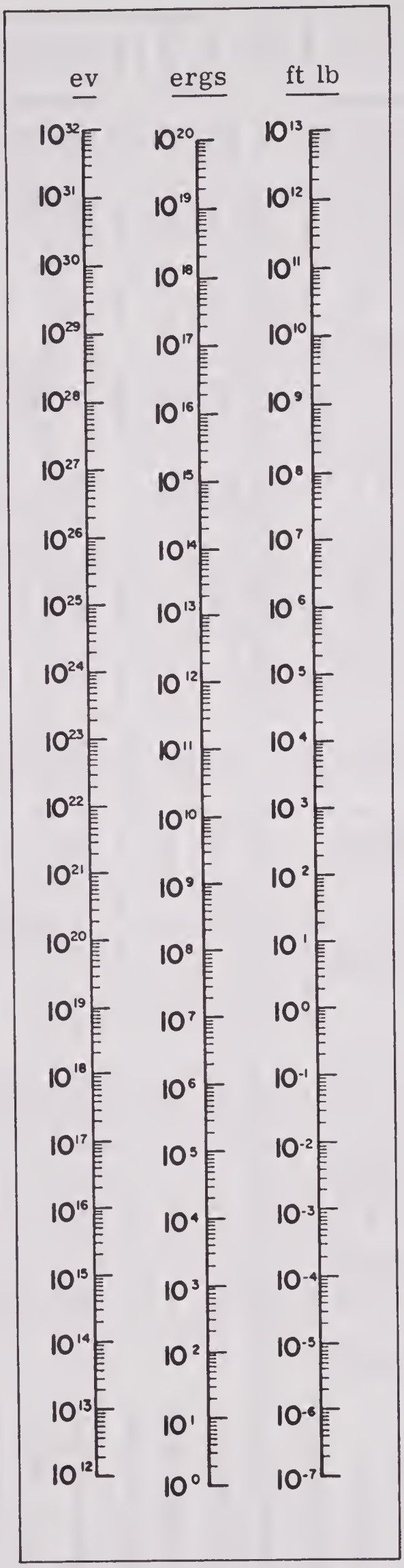


FIG. 22-5. Energy Conversions [3]

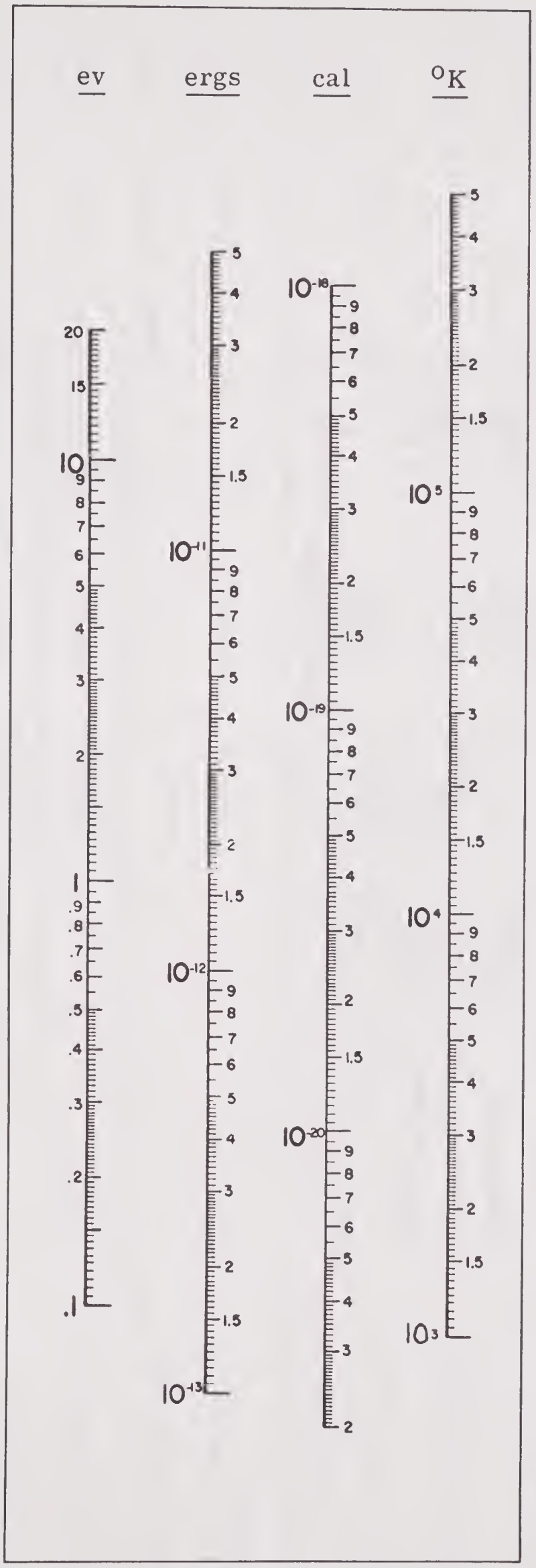


FIG. 22-5 (*Continued*). Energy Conversions [3]

TABLE 22-16. SPECTROSCOPIC ENERGY CONVERSIONS AND EQUIVALENCES [5]

		erg	electron volt	cycle per sec	wave number
1 erg	=	1	6.2419×10^{11}	1.509×10^{26}	5.035×10^{15}
1 electron volt	=	1.60207×10^{-12}	1	2.418×10^{14}	8.066×10^{3}
1 cycle per second	=	6.625×10^{-27}	4.135×10^{-15}	1	3.336×10^{-11}
1 wave number	=	1.986×10^{-16}	1.240×10^{-4}	2.998×10^{10} (velocity of light)	1

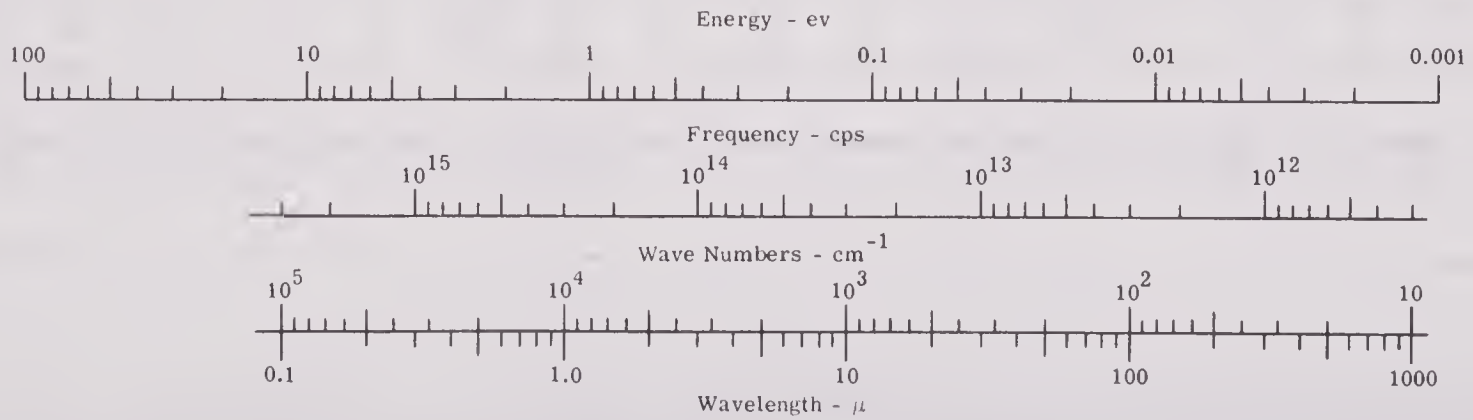


FIG. 22-6. Spectroscopic Energy Conversions

TABLE 22-17. POWER CONVERSIONS [2,4]

		w (abs)	w (internatl)	ft lb sec^{-1}	Btu min^{-1}	kg cal min^{-1}	erg sec^{-1}	hp	kg $m\ sec^{-1}$	g $cm\ sec^{-1}$
1 watt (absolute) (w)	=	1	0.999835	0.73756	0.056896	0.01433	1×10^{7}	0.0013410	0.101971	1.01971×10^{4}
1 watt (international) (w)	=	1.00165	1	0.73878	0.056990	0.01435	1.00165×10^{7}	0.0013432	0.10213	1.019878×10^{4}
1 British thermal unit (mean) per minute (Btu min^{-1})	=	17.580	1.355596	12.966	1	0.25192	1.7580×10^{8}	0.023575	0.138255	1.38254×10^{4}
1 foot pound per second (ft lb sec^{-1})	=	1.35582	17.577	1	0.077124	0.019433	1.35582×10^{7}	0.0018182	1.792661	17.92650×10^{4}
1 kilogram calorie (mean) per minute (kg cal min^{-1})	=	69.767	69.75548	51.457	3.969	1	6.9767×10^{8}	0.093557	7.114254	71.14210×10^{4}
1 erg per second (erg sec^{-1})	=	1×10^{-7}	0.999835×10^{-7}	7.3756×10^{-8}	5.688×10^{-9}	1.4333×10^{-9}	1	1.3412×10^{-10}	1.01971×10^{-8}	1.0971×10^{-3}
1 horsepower (U.S.) (hp)	=	745.70	745.5769	550	42.418	10.688	7.457×10^{9}	1	76.039774	760.3977
1 kilogram meter per second (kg m sec^{-1})	=	9.80665	9.80503	7.23299	0.55796	0.14053	9.80665×10^{7}	0.01315	1	9.999939
1 gram centimeter per second (g cm sec^{-1})	=	9.80665×10^{-5}	9.80503×10^{-5}	7.23299×10^{-5}	5.5796×10^{-6}	1.4053×10^{-6}	9.80665×10^{2}	1.315×10^{-7}	0.999993×10^{-5}	1

TABLE 22-18. ELECTRICAL UNIT CONVERSIONS [4-6]

	Electrostatic	*MKS*	*Electromagnetic* (cgs)
charge	2.9979×10^9 statcoulombs	1 coulomb	0.1 abcoulomb
electron charge	4.8029 statcoulombs	1.6021×10^{-19} coulomb	1.6021×10^{-20} abcoulomb
current	2.9979×10^9 statamperes	1 ampere	0.1 abampere
potential	3.3356×10^{-3} statvolts	1 volt	10^8 abvolts
capacity	8.9878×10^{11} statfarads	1 farad	—
electric field E	1 statvolt cm^{-1}	29979 volts $meter^{-1}$	—
electric displacement D	1 statvolt cm^{-1}	3.3356 coulombs $meter^{-2}$	—
dielectric constant of free space ϵ_0	1	0.11126×10^{-9} farad $meter^{-1}$	—
magnetic induction B	—	10^{-4} weber $meter^{-2}$	1 gauss
magnetic field H	—	10^3 ampere turns $meter^{-1}$	1 oersted
magnetic permeability of free space μ_0	—	10^{-7} henry $meter^{-1}$	1
resistance	1.1126×10^{-12} statohm	1 ohm	10^9 abohms
inductance	1.1126×10^{-12} stathenry	1 henry	10^9 abhenry

TABLE 22-19. PREFIXES

Multiples and Submultiples	*Prefixes*	*Symbols*
10	deca	da
10^2	hecto	h
10^3	kilo	k
10^4	10 kilo	
10^5	100 kilo	
10^6	mega	M
10^7	10 mega	
10^8	100 mega	
10^9	giga	G
10^{10}	10 giga	
10^{11}	100 giga	
10^{12}	tera	T
10^{13}	10 tera	
10^{14}	100 tera	
10^{-1}	deci	d
10^{-2}	centi	c
10^{-3}	milli	m
10^{-4}	100 micro	
10^{-5}	10 micro	
10^{-6}	micro	μ
10^{-7}	100 nano	
10^{-8}	10 nano	
10^{-9}	nano	n
10^{-10}	100 pico	
10^{-11}	10 pico	
10^{-12}	pico	p

Index

Abbé prism, 308
Abbé sine condition, 282, 428
Abbé V number, 403, 409
Aberrations
  Abbé V number, 403, 409
  afocal systems, 406
  angular, 381
  astigmatism, 384, 389, 400, 439, 448
  axial, 384, 389, 400, 404
  Bouwers system, 446
  catadioptric system, 448
  chromatic variations, 386
  correction, 385
  lateral color, 385, 390
  longitudinal, 384, 389, 400, 404
  off-axis, 385, 390
  refracting lens, 436
  stop shift theory, 404
  transverse, 385, 390
 coma
  calculation, 400
  catadioptric system, 448
  description, 382
  ray tracing, 388
  rim ray curve, 387
  spherical reflector, 438, 439
  stop shift theory, 404
  transfer function, 629, 634
correction, 385, 406, 570
distortion
  calculation, 400
  description, 389
  shock wave, 842
  stop shift theory, 404
  transfer function, 630
  window induced, 826
field curvature, 384
higher orders, 399
image curvature, 630
longitudinal, 381, 386
Petzval surface, 384, 400, 439, 440
plane parallel plate, 447
plots, 399
residual, 386, 407
Seidel, 382, 630
spherical
  blur spot, 435
  Bouwers system, 446
  calculations, 400
  catadioptric system, 447
  description, 382, 388
  fifth-order, 633
  longitudinal, 386, 399
  resolution limits, 414
  rim ray curve, 387
  refracting lens, 436
  spherical reflector, 438
  transfer function, 629, 631
  transverse, 399
 stop shift theory, 404
 symmetrical principle, 406
 third-order, 399, 402, 404
 transverse, 381, 399
 variations with aperture and image size, 385
 wave-front, 616, 630
 wave theory, 630
 zonal, 386
Aberration-free systems, 629
Abney grating mounting, 312
Absolute stability, *see* Stability, absolute
Absorptance
 approximations to band models
  nonoverlapping case, 201
  strong-line, 200
  table, 199
  weak-line, 197
definition, 784
directional, 26
Doppler coefficient, 196
Doppler line shape, 190
Doppler broadening, 191
Elsasser model
  compared to statistical model, 195
  development, 192
  strong-line approximation, 200
  weak-line approximation, 198
  weak-line model, 197
Lorentz broadening, 191
Lorentz coefficient, 196
Mayer-Goody model
  approximations of band absorptance, 199
  development, 194
models, table, 199
quasirandom model
  development, 196
  strong-line approximation, 201
  weak-line approximation, 199
random Elsasser model
  development, 196
  strong-line approximation, 201
  weak-line approximation, 199
single spectral line with Lorentz shape, 190

solar
ultraviolet, 792
various materials, 800
square coefficient, 196
square root region, 191
statistical model
approximation of band absorption, 199
comparison with Elsasser model, 195
strong-line approximation, 191
total, 28
Absorption
See also Absorptance
atmospheric bands, 238
nonequilibrium conditions, 28
selective filters, 286
wavelengths of molecular groups, 325
Absorption coefficient, 189, 190, 358
Absorptivity
definition, 784
measurement techniques, 796
ratio to emissivity, 787, 791, 804
selected materials, 805
Accelerated life test, detectors, 461
Acceleration
conversion factors, 870
definition, 856
Acetylene smoke, 356
Achromats, 282, 409
Acquisition, 744-748
Admiralty Research Laboratory blackbody slide rule, 17
Aerial targets, *see* Targets
Aerodynamic conditions, 827
Aerodynamic flow, optical influences, 826
Aerodynamic heat transfer, 827, 838
Aerographic film, 571, 576
Aerosols, 141, 208
Afocal system, 406, 422
Afterburners, 60
Air, 839, 851
See also Atmosphere
Air mass, 97
Air-to-air intercept, 744
Airglow, 104
Airy disc, 410, 440
Airy pattern, 616
Airy system, 623
Albedo
cylindrical geometry, 135
definition, 785, 812
hemispheric geometry, 814
planar geometry, 816
spectral distribution, 790
Alumina, 359, 804
Aluminum
α/ϵ ratio, 804
emissivity, 359
reflectance, 351
Aluminum foil, 77
Aluminum paint, 78, 809
Aluminum-oxygen group, 325
Alzac, 354
AM-FM discriminator, 745
Amici prism, 308
Amplifiers
See also Preamplifiers
feedback, 686
grounded cathode, 595
high-frequency response, 595
low-noise, 597
Miller effect, 595
noise sources, 584
power supply, 597
transistors
bias point, 598
bias stabilization, 598
circuits, 602
field-effect, 606
junction, shot noise, 586
low-noise, 606
noise, 597, 599, 602
planar, 608
silicon, 601
stability, 598
transformer coupling, 604
vacuum tube, 592
very high impedance, 596
Amplitude, complex, 615
Anamorphic systems, 424
Angle
conversion factors, 865
definition, 855
Angular aberrations, 381
Angular dependence of reflectivity, 793
Anode current fluctuations, 585
Antimony, 805
Apertures, *see* Detectors; Filters; Optics
Aplanatic lenses, 380, 427, 434
Aplanatic surfaces, 427
Apodization, 636
Luneberg theorem, 621
Area, 855
Argon, 179
Arsenic-modified selenium glass, 321
Arsenic trisulfide, 282, 294
Artificial sources, 32
Asphalt, 83, 89
Aspheric surfaces, 385, 402
Astigmatism, 384-389, 400, 439, 448
Astronomical data, 794
Astronomical telescope, 422
Atmosphere
See also Backgrounds, sky; Stratosphere
absorption, 178, 189, 237, 266
carbon dioxide, 180, 238
carbon monoxide, 249
composition, 179
density, 177
Lorentz broadening, 178, 189
methane, 250
nitrogen, 178
nitrous oxide, 246
oxygen, 178
ozone, 185
particles, 187

pressure, 189
radiation, 97, 737
scattering, 203
scintillation, 210
See also Stars, stellar scintillation
solar spectrum measurements, 227
spatial filtering, 737
temperature, 96, 177
transmission, 252, 737
water vapor, 181, 244
Aurora, 101
Autocorrelation function, 649, 726
Azimuth coverage, 738
Axial chromatic aberration, 384, 404
Axial cone, 377
Axis of symmetry, 844

Back focus, 375
Backgrounds
See also Terrain
discrimination, 732, 744
earth, 115
fluctuations, 730
marine, 166
radiation, 737-739, 752, 826
radiation-noise-limited system, 752
sky, 96, 143
See also Clouds
stellar radiation, 107
temperature, 462
Background-limited D^*-Q trade-off, 753
Baffles, 379, 442
Baird-Atomic filters, 299
Bands, *see* Spectral bands and lines
Bandpass interference filters, 290
Bandwidth, spectral, 763
Bang-bang systems, 715
Barium fluoride, 294
Barnes Engineering Co. catadioptric systems, 285, 286
Barnes Engineering Co. nitrogen pressure generator, 530
Barr and Stroud glass, 321
Base width (BW) filters, 287
Basic period filters, 290
Bausch and Lomb filters, 299
Bausch and Lomb glass, 321
Beryllium, 805
Bessel function, 648
Beutler radius grating mounting, 312
Bias, 502
Bias point, transistors, 598
Bilayer, fictitious, 292
Bilinear transformation analysis, 702
Birefringent crystals, 298
Blacks, 359-364
Black enamel, emissivity, 359
Blackbodies
cavity sources, 51
curves, 17
detectivity, 466, 467
emittance, detectors, 501
gold point, 44
Kirchhoff's law, 9
NBS standards, 38
noise equivalent power, detectors, 466
Planck's function, 795
quantum rates, 10
radiance, 96
reference, 761
responsivity, detectors, 464
simulator, 32
slide rules, 11
tables, 21
2400°K, 47
vertical, 45
Blackened chopper, 760
Blazed diffraction gratings, 310
Block blackbody slide rule, 11
Blocking filters, 291
Blur spot, 435, 452
Bode diagrams, 667
Bode method of linear systems analysis, 668, 678
Boil, 210
Bolometric detectors, 459, 497, 498
Boron nitride, emissivity, 359
Boundary-layer flow, 170, 828, 831, 846, 848
Bouwers-Maksutov optical system, 285, 445
Boxcar circuits, 696
Brick, daytime radiance, 146
Broadband transformer, 604

Cadmium, 805
Cadmium-germanium detectors, 491
Calcium aluminate glass, 317, 321
Calcium fluoride
prisms, 309
reflection coefficient, 298
substrate, 294
Canonic form, differential equations, 704
Caps, auroral, 101
Capacitance impact effects, 592
Carbon arc source, 49
Carbon-chlorine group, absorption wavelengths, 325
Carbon dioxide
atmospheric absorption bands, 238, 239, 240, 242, 243
atmospheric composition, 179, 180, 181
exchange cycle, 180
Carbon-hydrogen group, absorption wavelengths, 325
Carbon-oxygen group, absorption wavelength, 325
Carbon filaments, 38
Carbon monoxide
atmospheric absorption, 249
atmospheric distribution, 179, 187
Carbonyl group, absorption wavelength, 325
Cascaded optical systems, 626
Cascaded thermoelectric cooling systems, 525
Cascode connection, preamplifiers, 608
Cassegrain radiometers, 762
Cassegrain telescope, 442
Catadioptric systems, 285, 443, 624
Cavities, 32, 51

Celestial sphere, 785
Cell bias, detectors, 502
Cesium bromide prisms, 297, 309
Cesium dioxide film, 295
Cesium fluoride film, 295
Cesium iodide, prisms, 309
Childs's law, 585
Chiolite film, 295
Choppers, 759, 760
Chopping system, 759
Christiansen filters, 295
Christiansen wavelengths, 296
Chromatic aberration, 384-390, 400, 404, 436, 446-448
Chromium, 805
Circular aperture filters, 654, 655
Circular apertures, 616-618, 624, 630
Circular scan, 736
Circular-sectored reticles, 653, 655
Cirrus clouds, 99, 122
Clamping circuits, 696
Claude cycle analysis program, 545
Clock plot, aberrations, 388
Closed-loop describing function, 708
Closed-loop transfer function, 722
Clouds
  attitudes, 118
  cirrus, 99, 122
  cover, 119, 124, 737
  infrared transmission, 258
  meteorology, 118
  nacreous, 124
  noctilucent, 124
  stratospheric, 124
  top, 122
Coal tar pitch, reflectance of, 84
Coarse grain film, 572
Cobalt, 805
Codit silver paint, 90
Coddington's equations, 392-395
Coherent illumination, 626
Cold junction, heat balance at, 550
Cold trap, 533
Color reversal film, 571
Colored glass filters, 306
Columbium, 805
Collimator, 440
Coma, 400-404, 629-634, 382-389
Commutative condition, 614
Compensation
  high-frequency, 590
  image motion, 755
  singular points, 723
Complex amplitude, 615
Computers, 394, 640
Concave diffraction gratings, 310-311
Concrete, 88-92, 149-156
Condensers, optical, 424-430
Conductance, 786
Conduction
  definition, 785
  thermal joint, 786
Conductivity, thermal, of optical materials, 331
Cone, aperture, 377
Cone condensers, 427, 430
Conical lenses, 843
Conjugate functions, 647
Conrady G sums, 341, 403
Constant deviation prism, 308
Contact conductance, 786
Contours
  s-plane, 673
  sensitivity, 510
Contrast ratio, 623
Control systems
  asymptotic stability, 704
  block diagram, 667
  design, 718, 727
  feedback, 662
  integral control, 721
  linear systems, 662
  nonlinear systems, 703
  ramp inputs, 672, 687, 690, 694
  sampled-data system, 695
Conversion factors
  acceleration, 870
  angles, 865
  density, 867
  electrical units, 880
  energy, 855, 858, 875
  force, 871
  length, 860
  mass, 866
  power, 879
  pressure, 872
  ranges, 863
  spectroscopic energy, 878
  temperature, 873
  time, 868
  torque, 871
  velocity, 869
  volume, 864
  work, 875
Convolutional integral, 621, 648
Coolers, *See* Detectors, cooling systems
Copper
  Cu-Cu_2O detectors, 472
  Ge:Cu detectors, 490
  reflectance, 351
  thermal radiation properties, 805
Corner cube mirror, 773
Corning glass, 317, 318, 322
Correlation quality, 638
Corrugated metal, 93
Cotton, reflectance of, 76
Coupling, transistor-transformer, 604
Cross-correlation function, 726
Crossover frequency, 215, 664
Cryogenic data, 521
Cryolite, 295
Cryostat coolers, 523, 532
Crystals, birefringent, 298
Cube mirror, corner, 472
Curvature
  field, 384-389, 404
  image, 630

Petzval, 385, 439-440
Cutoff slope, filter, 287
Cutoff wavelength, filter, 287, 468
Cuton slope, filter, 287
Cuton wavelength, filter, 287

*D**, *see* Detectivity
*D**-*Q* tradeoff, 753
*D***, *see* Detectivity
Data enumeration, 460
Data function, equivalent continuous, 697
Day airglow, 106
Decimal equivalents of common fractions, 862
Decoy discrimination, 744
Deep space probes, 811
Defect, fidelity, 637
Definition, optics, 410
Defocusing coefficient, 387, 618, 630, 632
Degrees, radian equivalents, 865
Delta function, Dirac, 648, 662
Demodulation process, 745
Density
  definition, 856
  conversion factors, 867
  flux, 730
  photographic film, 573
Depth of field, 380
Depth of focus, 380
Describing function, 706-708
Detection
  definition, 738
  probability, 738, 740
  range, 739
  synchronous, 610
Detective quantum efficiency, 462
Detectivity
  blackbody, 466
  cutoff wavelength, 468
  blackbody *D**, 467
  detective quantum efficiency, 468
  maximized *D**, 467
  peak wavelength, 467
  photon-noise derivation, 513
  spectral, 466
  spectral *D**, 467
  spectral *D***, 467
  system design considerations, 515
  theoretical limit, 512
Detector-noise-limited system, 751
Detectors
  See also Detectivity
  accelerated life test, 461
  acceleration specifications, 461
  apertures, 459
  array, 755
  cooling systems
    design criteria, 522
    dewars, 559
    direct contact, 523, 526
    expansion engine cooling systems, 523, 540-545
    Joule-Thomson (cryostat), 523, 532
    limitations of fluid cooling, 520
    temperature ranges, 525
    thermoelectric (Peltier) coolers, 523, 546
  cruciform, 747
  Cu-Cu_2O, 472
  data enumeration, 460
  dewar flask, 459
  front-end description, 747
  GaAs, 471
  Ge:AuSb (*n*-type), 488
  Ge:AuSb (*p*-type), 489
  Ge:Cd, 491
  Ge:Cu, 490
  Ge:Hg, 492
  Ge-Si:Au, 496
  Ge-Si:Zn, 495
  Ge:Zn, 493
  Golay, 772
  HgTe, 494
  InAs, 481, 482
  InSb, 484-487
  InSb bolometers, 497
  PbS, 474-476
  PbSe, 477-479
  PbTe, 480, 836
  noise, 464, 588
  parameters, 459, 462
  photodetectors, 458
  radiation transducer, 458
  saturation, 833
  Si, 473
  symbols and units, 459
  system, multiple, 753
  Te, 483
  test procedures
    frequency response, 504
    NEP, determination of, 501
    noise spectrum, 510
    optimum bias, 502
    pulse response, 505
    sensitivity contours, 510
    spectral response, 508
    spinning mirror technique, 506
    time constant, 504
  thermal
    Golay cells, 458, 459, 500
    thermistors (bolometric), 458, 497, 498
    thermocouples (thermovoltaic), 458, 459, 499
    thermopiles (thermovoltaic), 458, 459
  time constant, 504
  vacuum environment, 461
  windows, 359
Deuterium, 520
Deviating prisms, 308
Dew point, 184
Dewars, 459, 557
Dielectric constants, optical materials, 329
Differential equation, canonic form, 704
Diffraction, *See* Optics, diffraction theory
Diffraction integral, Kirchhoff, 635
Diffraction gratings, 309-312, 763-772
Diffraction limitation of optical system, 440
Diffraction-limited systems, 448

Diffraction pattern, 411, 635
Dimensionless wave number, 288
Diode, space-change-limited, 586
Dirac delta function, 648
Direct-contact cooling systems, 522, 526
Direct-ratio spectrophotometer, 776
Directional absorptance, 26
Directional emissivity, 23, 26
Directional emittance, 785
Directional reflectance, 23, 26
Directional reflectivity, 88-93, 793
Discrimination, 732, 737
  decoys, 744
Discriminator, AM and FM, 745
Disc, Airy, 410, 440
Dispersing elements, 763
Dispersing prisms, 309
Dispersion
  angular, of thin lenses, 340, 404
  Herzberger equation, 337
  optical materials, 328
    equation, 335
    power, 339
    refractive index, 338
  partial, 403
  reciprocal, 403
Displacement error, systems, 663
Displacement input, linear systems, 689
Displacement law, Wien, 10
Displacer piston expander, 544, 545
Distance measures, 863
Distortion, 384-389, 400-404, 630, 826, 842
Distribution
  albedo radiation, 790
  atmospheric constituents, 181
  illuminance, 624
  innerloop gain, 725
  Poisson, 585
  spectral, 761
  stellar radiation, 110
Diurnal variations
  atmospheric scintillation, 223, 226
  radiance of concrete, 149
  radiance of grass, 143
  radiance of backgrounds, 144
Domain
  frequency, 637
  spatial, 623
    of linear systems, 614
  spatial frequency, 628
Doppler coefficient, 196
Doppler effect on Lorentz broadening of spectral lines, 189
Doppler half-width, 190
Doppler line shape, 190
Double monochromators, 765
Double-beam optical null spectrophotometer, 775
Double-beam photometer, 772
Double-pass spectrometer, 771
Double-pass system, monochromator, 773
Dove prism, 308
Drag penalty, 827
Drift, radiometer, 759
Droplets, atmospheric, 188
Dupont Flat Black paint, emittance, 356
Dwell time, 730, 734, 739, 752
Dyed-plastic filters, 307
Dynamic transconductance, 586
Dynasil glass, 318

Eagle grating mounting, 312
Earth
  as a background, 115
  astronomical data, 794
  emissive power curve, 790
  shine, 785
  spectral emissive power curve, 790
Eastman Kodak Co.
  film and plates, 571-580
  filters, 299, 575, 580
  lenses, 282
  NOD-18 black, 359
Ebanol, 805
Ebert-Fastie plane grating mountings, 311
Ecliptic, 785
Effective optical thickness, 294
Effective power, radiometers, 763
Ektachrome, Aero film, 571
Eikonal (wave-front aberration), 616
Electrical units
  conversion factors, 880
  definition, 857
Elevation angle, ocean spectral radiance, 169
Elsasser model, 192-200
Emissance, *See* Emissivity
Emission
  atmospheric constituents, 96
  exhaust gases, 67
  ozone, 96
  sky, 96
  targets, 70
Emissivity
  alumina, 804
  cavities, 32
  chopper mirror, 760
  definition, 784
  DeVos method, 33
  directional, 26
  Globars, 359
  Gouffé method, 32
  hemispheric, 793
  Kirchhoff's law, 9
  materials, 359, 800, 805
  measurement techniques, 796
  NBS comparison standards, 44
  nonequilibrium condition, 28
  normal, 795
  partially transparent bodies, 23, 349
  radiant, 74
  ratio to absorptivity, 787, 791, 804
  Sparrow method, 37
  terrain, 75, 142
  total, 28
  water, 167
  windows, 827, 835

Emittance
  blacks, 362, 356
  definition, 784
  radiant, 23
  total, 785
Emulsion, infrared, 570
Enamel
  black, 359
  Chinese red, 80
  white, 81, 359
Endless checkerboard, 651, 654-655
Energy
  conservation of, 27
  conversion factors, 855, 858, 875
  definition, 856
  distribution, 410
  Fermi, 544
Engelhard glass, 318
Entrance aperture, flux density, 730
Entrance pupil, 377, 379, 733
Entrance slit, monochromator, 766
Episcotister, 653
Epoxy-type molding compound, 357
Equatorial inclination, 785
Equilibrium constant, 66
Equivalent, Thevenin, 590
Equivalent continuous data functions, 697
Equivalent layer, Herpin, 293
Equivalent power detectors, 466
Equivalent noise voltage, 607
Equivalents of common fractions, decimal, 862
Erectors (optical systems), 424
Errors
  acceleration, linear systems, 663
  control methods, 721
  displacement, linear system, 663
  generation, tracking systems, 750
  modulation techniques, 744
  rate, linear systems, 663
Error plus error-rate system, 714
Ester group, absorption wavelength, 325
Eurasian land mass, 122
Exhausts, 59-67
Exit heating, rockets, 68
Exit pupils, 377
Expansion, thermal, of optical materials, 332
Expansion-engine cooling systems
  characteristics, 523
  commercially available, 541
  displacer piston expander, 544, 545
  Gifford-McMahon piston expander, 540
  turbine expanders, 544, 545
Exposure, films and plates, 573
Exposure meters, 570
Extinction coefficient, 351

f number, 379, 429
Fabrics, reflectance, 76
Fabry-Perot interferometer, 290, 778-781
Farrand Optical Co. filters, 299
Feedback, negative, 593
Feedback amplifiers, 686
Feedback control system, 662
Fermi energy, 544
FET preamplifier design, 608
Fidelity defect, 637
Field, depth of, 380
Field curvature, 384-389, 404
Field-effect transistors, 606
Field lens, 424-429
Field of flow refraction, 839
Field of view, instantaneous, mapping system, 734, 739, 745
Field stop, 378
Fifth-order aberrations, 382, 633
Figure-of-merit, cooling systems, 548
Films and plates
  aberration correction, 570
  aerographic, 571, 576
  characteristic curves, 573
  coarse grain, 572
  color reversal, 571
  density, 573
  Ektachrome Aerofilm, 571
  exposure, 573
  extremely high-contrast, 572
  H and D curves, 573
  high contrast, 571
  high-speed, 571, 575, 579, 580
  infrared, 571-580
  medium contrast, 572
  reciprocity characteristics, 574
  relative visibility curve, 573
  sensitometric characteristics, 573
  spectral sensitivity, 574
  spectroscopic, 572-579
Filtering, spatial, 646, 737
Filters,
  absorption, 306
  absentee layers, 288
  angle shift, 290
  background region, 287
  birefringent, 298
  blocking, 291
  characteristic admittance, 293
  Christiansen, 295
  circular aperture, 654-655
  circular-sectioned reticles, 655
  commercially available, 299
  dyed-plastic, 307
  film materials, 295
  fixed-field, moving reticle, 650
  Fourier transforms, 654
  infinite checkerboard reticle, 654-655
  infinite parallel-spoke reticle, 654
  interference
    analogies with transmission-line theory, 293
    angle of incidence, 293, 350
    bandpass, 290
      characteristic matrix, 291
    description, 236
    effective optical thickness, 294
    fictitious bilayer, 292
    films, 294

filter matrix, 291
Herpin equivalent layer, 293
long-wavepass, 290
narrowband, 291
polarization, 298
short-wavepass, 290
spike, 292
square-band, 291-292
substrates, 294
terminology, 286, 288
line admittance, 293
moving-reticle fixed field, 650
moving-reticle scanning-field, 652
narrow, radiometer use, 763
plastic, 306
rectangular aperture, 654-655
scanning-aperture space, 650
scanning-field, moving reticle, 652
selective reflection, 297
selective refraction, 298
sintered metal, 533
spatio-temporal output, 649
spectral, monochromators, 766
substrate materials, 294
terminology, 286
transmittances, 574, 580
First-derivative control methods, 721
First-order aberrations, 382
First-order optics, 371
First-order high-reflectance zone, 288
Fish Schurman Corp. filters, 299
Fixed field, moving-reticle space filters, 650
Flame-temperature spectrometer, 772
Flex-o-lite beaded paint, 92
Flicker current noise, 586
Flow, boundary-layer, *see* Laminar flow; Separated flow; Turbulent flow
Fluid cooling systems, 520
Fluorite, 39
Flux, meteoric, 793
Flux density, 730
FM-AM discrimination, 745
Focal length, effect of shock wave, 844
Focal plane, paraxial, 386
Focal point, 371
Focus
back, 375
depth of, 380
paraxial, 634
plane of best, 638
sagittal, 384
Fog, 188, 207, 211
Folded optical systems, 442
Force
conversion factors, 871
definition, 856
Fore prism, 766
Folded reflector, 441
Four-element achromat, 284
Fourier approach (cascaded system), 627
Fourier transforms, 614, 621-622, 646-652
Fractions, decimal equivalents, 862
Fraunhofer lines, 227
Fraunhofer plane, 616
Fraunhofer receiving plane, 619
Free filter range, 287, 290
Frequency
crossover, 215, 664
spatial, 626
Frequency domain, 637
Frequency filtering, 737
Frequency response, *see* Response, frequency
Frequency shifting, space, 646
Front-end description, 747
Frost point, 184
Fuels, high-energy solid, 67
Fused silicate glass, 317

G sums, Conrady, 341, 403
Gain
adjustment design methods, 718
distribution, inner-loop, 725
margin, 664
system, 676
Galactic concentration of stars, 109
Galilean telescope, 422
Gallium arsenide
detector, 471
reflection coefficient, 297
transmission, 324
Gallium antimonide
reflection coefficient, 297
transmission, 324
Gallium phosphide, transmission, 324
Gas-supply cooling systems, 532
Gases, liquefied, 520
Gaussian equation, 372
Gaussian optics, 371
Gaussian probability, 740
Gaussian quadrature method, 631, 633
General Electric Co.
blackbody slide rule, 11
glass, 318, 321
Generation-recombination noise, 470, 587
Geometry
cylindrical, of albedo, 815
hemispheric, albedo, 814
incident beam, 25
planar, albedo, 816
planetary thermal radiation, 817
problem, in systems design, 737
projected solid angle, 24
radiation, 22
scattering angle, 117
sea-surface, 168
Germanate glass, 321
Germanium
filter film, 295
filter use, 294
detectors
Ge:AuSb (*n*-type), 488
Ge:AuSb (*p*-type), 489
Ge:Cd, 491
Ge:Cu, 490
Ge:Hg, 491
Ge-Si:Au, 496

Ge-Si:Zn, 495
Ge:Zn, 493
lenses, 282
semiconductor properties, 324
Germanium monoxide, 352
Germanium-oxygen group, 325
Gifford-McMahon piston expander, 540
Gladstone-Dale constant, 839
Glare stops, 379
Glass
arsenic trisulfide, 294
calcium aluminate
index of refraction, 323
description, 317
transmission, 323
types, 321
colored filters, 306
Corning, No. 186SJ, 317, 318, 322
filter use, 294
fused silicate, 317
fused quartz, 317
General Electric Co., 321
germanate, 321, 322
high-silica-description, 317, 318
lead silicate, 318
lenses, 282
NBS F998, 321
nonoxide, 321
silicate, 321
suppliers, 318
thermal radiation properties, 805
Vycor, 317
Global stability, 704
Globar source, 48, 359
Golay cells, 458, 500
Golay detectors, 772
Gold
Ge:AuSb (*n*-type) detectors, 488
Ge:AuSb (*p*-type) detectors, 489
Ge-Si:Au detectors, 496
reflectance, 79, 351, 805
Gold black, 359
Gold paint, 79
Gold-point blackbodies, 44
Gouffé method of calculating emissivity, 32
Gradient methods of optical design, 409
Graphite, 806
Grass, daytime radiance, 146, 164, 157
Grating spectrometers, 766
Gratings, 309-312, 763-772
concave mountings, 311
Grazing incidence, 312
Great circle arc, 115
Gregorian telescope, 442
Greybodies, 70
Grounded-cathode amplifiers, 595
Grounding techniques, 611

H&D curves, 573
H&L layers, 288
Half-width
filters, 287
spectral lines, 189
Hardness, optical materials, 333
Heat balance at cold junction, 550
Hazes, 188, 207, 211
Heat
See also Temperature; headings beginning Thermal
flow, 786
flux, solar, 811
load, of cooling systems, 521
pumping, 550
spacecraft, 68
specific, 333
transfer
aerodynamic, 827
flat plate, 828
to hemisphere, 829
radiant, 787
radiation, 828
separated flow, 838
wall temperature, 829
transfer coefficient
definition, 827
flat plate, 828
wakes, 839
Heat-energy balance, 73
Helium, 179, 520
Helmholtz's reciprocity theorem, 24, 27
Hemispherical emissivity, 793
Hemispherical emittance, 785
Hemispherical heat transfer, 829
Hemispherical immersion lens, 433
Hemispherical integration, 25
Hemispherical shock-wave, 846
Heraeus glass, 318
Herpin equivalent, 290, 293
Herzberger dispersion equation, 337
High-contrast film, 571
High-frequency compensation, 590
High-frequency response amplifier, 595
High-intensity sources, 49
High-order aberrations, 399
High-silica glass, 318
High source impedance, 590
High-speed flight, 826
High-speed infrared film, 571, 575, 579, 580
High-reflectance zone filters, 288, 290
Holding circuits, 696
Homogeneous paths, 263
Horizon, 143, 737
Horizontal paths, 252, 263, 266
Humidity, 97
detector requirements, 461
Hydrogen, atmospheric, 179
Hydrogen deuteride, 520
Hydroxyl group, 325
Hypersensitizing, 570
Hypocycloid scan, 736

Ice crystals, 122
Ideal generator, noise power, 588
Illuminance
definition, 615
distribution, 624

ratio, 615
Illumination, coherent, 626
Image
blur, 386
spot size, 449
boil, 212
contrast, 643
curvature, 630
dancing, 849
enlargement (atmospheric scintillation) 213
motion compensation, 755
nutation, 747
size aberrations, 385
spot size, 410
Immersion lenses, 427, 433-4
Impedance
amplifier, very high, 596
detectors, 462
high source, 590
matching, 293
matching circuit, detectors, 502
open-circuit, filters, 293
short-circuit, filters, 293
source, 609
vs noise, 598
Impulse response, 614
In-line spectrometer, 774
Incidence, angle of, 293, 350
Incident beam geometry, 25
Incident radiance, 27
Incident radiation
albedo, 813
planetary, 819
solar, 813
total, space, 821
Inclination
equatorial, 785
planetary, 785
Incoherent light, 621
Inconel, 359, 806
Inconel-X, 800
Index of refraction
air, 839
boundary-layer effects, 847
calcium aluminate glass, 323
Corning No. 9752 glass, 322
dispersion values of optical materials, 338
NBS F998 glass, 321
optical materials, 328
shock-wave effect, 844
water, 167
Indium antimonide
bolometers, 497
detectors, 484-487
reflection coefficient, 297
semiconductor properties, 324
Indium arsenide
detector, 481, 482
reflection coefficient, 297
transmission, 324
Indium phosphide
reflection coefficient, 297
transmission, 324
Infrared systems design
general procedures, 730
gross analysis, 730
mapping systems, 750
optical systems, 732
scanning dynamics, 734
search system, 737
sensitivity calculation, 731
tracking system design, 743
Inner loop gain distribution, 725
Input
control, 721
displacement, 689
noise equivalent, 731, 740, 753
parabolic, 691
ramp, 672, 687, 690, 694
step, 672, 689, 693
Insolation, *see* Radiation, solar
Instantaneous fields of view, 739
Insulating materials, 521
Insulators, properties, 522
Integral, convolution, 621
Integral coolers, 529
Integration time, 745
Intercept, air-to-air, 744
Interchangeable gratings, 771
Interchangeable prisms, 721
Interference filters, 286-294
Interference fringes, 781
Interferometer-spectrometer, 774
Interferometers
Fabry-Perot, 290, 778, 781
spherical, 779
Lummer-Gehrcke plate, 779
Michelson, 777
Rayleigh, 776
spectral transmittance, 780
Twyman-Green, 777
Invariance, spatial, 621
Invariant, Lagrangian, 433
Inverse-Nyquist method, 677
Inverse *z*-transforms, 698
Inverting prism, 309
Iridium, 806
Iron, 806
Iron oxide
emittance, 362
paint, 87
Irradiance
celestial bodies, 112
definition, 784
NBS standards, 40
optical systems, 380
visual magnitude, 114
Irtran, 282, 294, 326
Isoplanatic region, 621
Isoplanatic systems, 614

James-Weiss stability criterion, 672
Jet propulsion systems, *see* Rocket and jet propulsion systems
Johnson noise, 469

Johnson solar spectral curve, 789
Joint, thermal, conduction at, 786
Joule-Thomson cooling systems, 523, 532
Junction, cold, 550
Junction transistor, shot noise, 586
Jupiter, 794

K-Monel, 804
Kel-F, 326
Kirchhoff diffraction integral, 635
Kirchhoff diffraction theory, 615
Kirchhoff's law, 9, 27, 72, 795
KRS-5 prisms, 309
Krylon flat black paint, 362, 809
Krypton, 179

Lacquer, 359
Lagrangian invariant, 373, 433
Lambertian sources, 22
Lambert's law, 70, 795
Laminar flow, 828
Lamps, 40, 41, 49, 643
Lamp black, 359
Lateral chromatic aberration, 390-400
Lateral magnification, 373
Layers
  filters
    absentee, 288
    H&L, 288
  Herpin equivalent, 293
  single, reflection and transmission, 348
Lead, 806
Lead selenide detectors, 477, 478, 479
Lead silicate glass, 318
Lead sulfide detectors, 474, 475, 476
Lead telluride
  detectors, 480, 836
  filter film, 295, 324
Leidenfrost transfer, 531
Length, conversion factors, 860
Least squares design, 408
Lenses
  achromats, 409, 282
  aplanatic, 380, 427, 434
  conical, 843
  Conrady G sums, 341
  corrector, 286
  design, 640
  field, 424, 732
  immersion, 427, 433
  materials, 282
  multielement, 282-284
  projection, 426
  relay, 426
  S.C.A, 282, 284
  single element, 282
  thin lens angular dispersion, 340
Liaponov stability analysis, 704, 710
Liaponov theorem, 711
Light, incoherent, 621
Light pipes, 427, 430, 431
Linear systems
  analysis methods, 668
  definitions, 662
  feedback control system, 662
  general concepts, 614
  isoplanatic, 614
  operator, 614
  ramp inputs, 672, 687, 690, 694
  spatial domain, 614
  stationary (time invariant), 614
  superposition integral, 614
  type 0, 686
  type 1, 688
  type 2, 692
  transfer function, 664
Linearization methods of optical design, 409
Lines, spectral, 194, 291-293
Linfoot quality factors, 637
Liquid-feed coolers, 529
Liquid-supply cooling systems, 533
Liquid propellants, 64
Liquid transfer coolers, 526
Lithium fluoride
  Christiansen wavelengths, 296
  prisms, 309
  reflection coefficient, 298
  filter substrate, 294
Littrow double-pass system, 773
Littrow mirror, 771
Littrow monochromator, 764
Littrow-mounted spectrometer, 766
Littrow prism, 308
Loading resistance noise factor, 589
Log modulus, 668, 682
Long-wavepass interference filters, 290
Longitudinal aberration, 381-386, 399
Longitudinal chromatic aberration, 384-389, 400
Longitudinal magnification, 373
Look, direction of, 97
Loops
  closed, 708, 710, 722
  inner, gain distribution, 725
  motor driven, 688
  multiple, analysis, 725
  open, 707
  torqued, gearless, 692
  track, 692, 744, 746, 747
Lorentz broadening, 178, 189, 191
Lorentz line shape, 190, 194, 291
Lorentz-Lorentz law, 839
Loss tangent, 350
Low-noise amplifiers, 597
Low-noise cable, 611
Low-noise transistors, 606
Low-temperature liquids, 520
Lucalox glass, 321
Lummer-Gehrcke plate interferometers, 779
Lumped-constant line, 293
Luneberg apodization theorem, 621, 636
Lyot filter, 299
Lyot-Ohman filter, 298

M-Circle, 676
Mach number
  effect on shock lens, 843
  free-stream, 848
  relation to Stanton number, 829
  shock-wave angle, 840
Magnesia, 359
Magnesium, 806
Magnesium fluoride, 295, 836
Magnesium oxide, 294, 296, 362
Magnification, 373, 423
Magnifying power, 373, 423
Magnitude
  log, 678, 682, 720
  stellar, 107
Maksutov optical systems, 285, 445
Mangin mirror, 444
Manganin, 807
Mapping system, 734, 750-756
Marginal rays, 399
Margin
  gain, 664
  phase, 664
Marine backgrounds, 166
Mars, 794
Masonry, 82
Mass
  conversion factors, 866
  definition, 855
  units, table, 867
Matrix, filter, 291
Maximized D^*, 467
Maximum difference expression, 10
Maxwell's probability law, 585
Mayer-Goody model, 194
Mean-square noise voltage, 591
Medium contrast film, 572
Melting temperature, optical materials, 330
Meniscus lenses, 282
Meniscus corrector, 286
Mercuric telluride detectors, 494
Mercury (planet), 794
Mercury
  arc, 49, 50
  Ge:Hg detectors, 492
  HgTe detectors, 494
Merit factors, 637
Merit functions, 408
Meteoric flux, 793
Meteoroid bombardment, 792
Meteorology, 118
Methane
  atmospheric absorption bands, 250
  atmospheric, 179, 187
  solid, temperature, 562
Methyl group, 325
Metric system prefixes, 857, 881
Michelson interferometer, 777
Microphonics, 461, 592
Microscopic irregularities, 25
Mie scattering, 205, 206
Miller effect, 595, 608
Minutes, radian equivalents, 865
Mirrors
  chopper, 760
  commercially available, 285
  corner cube, 773
  Littrow, 771
  Mangin, 444, 452
  paraboloid, 773
  spinning, 506
Mixing ratio, 182, 183
Modulation noise, detectors, 469
Molding compound, epoxy-type, 357
Molecular emission of exhausts, 67
Molecular groups, 325
Molybdenum, 801
Monel, 807
Monochromators
  double, 765
  grating, 772
  Littrow, 764
  prism, 772, 766
  rapid-scan, 773
  single-pass, 765
  spectrometer use, 771
Moon, 794
Motor-driven loops, classical, 688
Moving-reticle filters, 650
Multicouple single-stage cooling systems, 555
Multielement lens, 282, 284
Multistage cooling systems, 555
Multiple-detector array, 755
Multiple-detector system, 753
Multiple-loop analysis and synthesis, 725
Mylar, 804

n-type detectors, 488
Nacreous clouds, 124
Narrowband filters, 291
National Bureau of Standards
  blackbody standards, 38, 44, 45
  carbon filament standards, 38
  comparison standards, 44
  glass, 321
  radiation standard, 39
  spectral irradiance standards, 40
  spectral radiance standards, 41
  total radiation standards, 38
Negative feedback, 593
Negative film, infrared, 571, 576, 580
Negative transfer function, 629
Neon, 179, 520
Neptune, 794
Nernst glower, 48
Networks, 718, 724
Newtonian optical equations, 372
Newtonian telescope, 441
Nichrome, 807
Nickel, 807
Night glow, 104
Niobium, 805
Nitrogen, 179, 520
Nitrogen-hydrogen group, 325
Nitrous oxide, atmospheric, 179, 187, 246
Noctilucent clouds, 124

Noise
amplifier sources, 584
anode current flucuations, 585
broadband figure, 609
factors, 587
detector, 588
loading resistance, 589
minimum, 602
operating, 588
overall, 590
preamplifier, 589
figures, 587, 606, 607, 609
flicker current, 586
generation-recombination, 470, 587
generator, 589
Johnson, 469
low
cable, 611
design, 605
modulation, 469
Nyquist, 469
output voltage, 464, 599
peaked-channel, 591
photon, 513
planar transistors, 603
power, ideal generator, 588
semiconductors, junction transistor, 586
shot, 470, 585, 586, 589, 591, 592, 597
signal-to-noise ratio, 594, 740, 748
source, 595
spectral, 466
spectrum, of detectors, 510
temperature, 469, 588
temperature-caused, detectors, 469
thermal, 469, 584, 591
total output, 591
transistor, 597-602
voltage, mean-square, 591
Noise equivalent input, 731, 740, 753
Noise equivalent power, 466, 501, 731, 753
Noise equivalent resistance, 589
Noise equivalent resistor, 590
Noise equivalent voltage, 607
Noise-limited systems, 751, 752
Nonlinear systems
analysis methods, 705
bang-bang system, 715
definitions, 703
error plus error rate system, 714
saturated, 713
Nonoverlapping line approximation, 201
Nonselective scattering, 206
Normal emissivity, 795
Normal emittance, 785
Normalization methods, 761
Nutation, 747
Nuvistor, 590
Nyquist methods
inverse, 677
linear system analysis, 668, 672
sampled-data systems, 702
Nyquist noise, detectors, 469
Nyquist plot, design method, 718
Ocean, 168-170
Off-axis chromatic aberration, 385
Off-axis response, 626
One-third rule, 725
Open-circuit impedance, 293
Open-loop describing function, 707
Operating noise factor, 588
Operating temperatures, detectors, 463
Optical Coating Laboratory filters, 299
Optical component, 371
Optical design
automatic, 408
electronic computers, 394
primary aberration correction, 406
general considerations, 406
merit functions, 408
techniques, 406
Optical difference (wave front aberration), 616
Optical element, 371
Optical materials
angular dispersion of thin lens, 340
blacks
emittance, 362
optical properties, 356
cooled
description, 359
transmittance, 361
Conrady G sums, 341
dielectric constants,
dispersion, 328
dispersion equations, 335
dispersive power, 339
equations, 347
extinction coefficient, 351
hardness, 333
Herzberger dispersion equation, 337
index of refraction, 328, 338
loss tangent, 350
radiation damage, 358
softening or melting temperature, 330
solubility, 333
specific heat, 333
surface coatings, 351
thermal conductivity, 331
thermal expansion, 332
transmittance, 364
Young's modulus, 333
Optical member, 371
Optical parameters, 742
Optical paths, shock wave, 843
Optical systems
See also Lenses
aberration-free, 629
afocal, 422
anamorphic, 424
aplanatic, 427
arrays, 755
cardinal points, 371-376
cascaded, 626
catadioptic objective
Bouwers (Maksutov) system, 285, 445
Maksutov, 285, 445
Mangin mirror, 444

Schmidt system, 443
transfer function, 624
condensers, 424
coverage, 732
detector systems
Abbé sine conditions, 428
f/no., 429
field lenses, 427, 429
immersion lenses, 427, 433
light pipes, 427, 430
vignetting, 429
diffraction-limited, 448
erectors, 424
image blur size, 450
merit factors, 637
multielement, 375
periscopes, 424
projection, 424
radiometers, 762
reflecting objectives
baffling folded, 442
Cassegrainian objective, 442
folded reflector, 441
Gregorian telescope, 442
Newtonian telescope, 441
optical collimator, 440
parabolic reflector, 440
spherical reflector, 437
refracting objectives, 435, 454
relay systems, 424
transfer functions
aberrations, 629
annular and annulus apertures, 624
apodization, 636
cascaded systems, 626
circular-aperture, 624, 630, 632
coherent illumination, 626
coma, 634
computation, example, 634
contrast ratio, 623
defocusing, 630
description, 621
expressions, 624
linear systems development, 614
modulus, 623, 633
negative, 629
sampled-data systems, 700
slit aperture, 630
spherical aberration, 631
Strehl criterion, 623
typical reflecting system, 624
windows, *see* Windows
Optical thickness, 288, 294
Optics
See also Lenses; Optical systems
Airy disc, 410
apertures
aberrations, 385
annular and annulus, 616, 624
circular, 616, 618, 624, 630, 654, 655
cone, 377
entrance, flux density, 730
limit, 626
rectangular, 619, 654, 655
slit, 617, 619, 630
stop, 371, 377-379
aplanatic systems, 380
back focal length, 376
baffles, 379
cardinal points, 371-376
combination of two elements, 376
definitions, 410
depth of field, 380
depth of focus, 380
diffraction theory
annular and annulus aperture, 616, 624
circular aperture, 616, 618
complex amplitude, 615
defocusing, 618
illuminance, 615
Kirchhoff theory, 615
point spread function, 615
pupil function, 616
rectangular and slit aperture, 619
Straubel pupil function, 618
Strehl criterion, 621
variable pupil functions, 620
effective focal length, 376
energy distribution, 410
entrance pupil, 379
equations, 372-375
f/no., 379
field stop, 378
focal points, 371
frequency response *see* Response, frequency
front-end description, 747
geometry, 371
glare stops, 379
image irradiance, 380
image spot size, 410
Lagrangian invariant, 373
magnification, 323
numerical aperture, 379
paraxial rays, 382
paraxial ray trace, 377
principal points, 371-375
principal ray, 376-377
pupils, 377
ray tracing, 375
Rayleigh criterion, 412
relative aperture, 379
resolution, 410
sign convention, 371
speed, 379
Snell's law, 371
symbols, 371
vignetting, 378
Optics Technology, Inc., filters, 299
Optimum bias of detectors, 502
Optimum noise figure, 607
Ordnance materials, 74
Oscillations (high-reflectance zone), 290
Output
control, first-derivative method, 721
noise voltage, 599
spatio-temporal, filters, 649

Overshoot, linear system, 662
Oxygen, 179, 520
Ozone, atmospheric
  distribution, 185
  emission peak, 96
  high-altitude maximum concentration, 187
  percentage, 179
  relation to air glow, 105
  seasonal variation, 186
  ultraviolet radiation, 185

p-type detectors, 489
Paint
  See also Enamel
  aluminum, 78, 809
  Centerlite white, 91
  Codit silver, 90
  Flex-o-lite beaded, 92
  gold, 79
  green masonry, 82
  Krylon black, 809
  No. 13, 86
  red iron oxide, 87
  thermal radiation properties, 807
Palladium, 809
Palmer scan, 736
Parabolic input, 691
Parabolic reflector, 440, 452
Paraboloid, optical properties, 440
Paraboloid mirror, 773
Parameters
  cooling, 742
  optical, 742
  sensitivity, 740
  signal processing, 742
  system block diagram, 742
  systems, 730
  windows, 836
Paraxial focal plane, 386
Paraxial focus, 634
Paraxial marginal ray, 402
Paraxial principal ray, 402
Paraxial rays, 377, 382, 390, 399
Parseval's theorem, 647
Partial dispersion, 403
Partial reflectance, 28
Particle-size distribution curves, 188
Paschen-Runge grating mounting, 312
Passband filters, 287
Path length, 116
Paths, atmospheric, 263-266
Peaked-channel noise, 591
Peltier coefficient, 548
Peltier coolers, 523, 546
Peltier couple, performance, 549
Pentode, 586, 592, 595
Periscopes, 424
Petzval contribution to aberration, 400
Petzval curvature, 385, 439, 440
Petzval surface, 384, 400, 439-440
Pfund grating spectrometers, 767
Pfund radiometer, 762
Phase angles, 736
Phase difference, filters, 288
Phase lag network, 718
Phase margin, linear systems, 664
Phase-plane
  analysis, 709
  definition, 704
  response time, 710
Phase plots (nonlinear systems), 720
Phase portrait, 709, 713
Photoconductive detectors, 458
Photodetectors, 458
Photoelectromagnetic detectors, 458
Photoemissive detectors, 458
Photographic film, *see* Films and plates
Photometer, double beam, 772
Photovoltaic detectors, 458
Photon emission rates, 106
Photon noise limitation, detectors, 513
Phthalocyanide, sublimated, 306
Physical constants, values, 855, 858
Pitch, coal tar, 84
Pitot pressure, 830
Planar transistors, 603
Planck blackbody function, 795
Planck's law, 9, 20
Plane of best focus, 638
Plane of the ecliptic, 785
Planetary albedo, 814
Planetary inclination, 785
Planetary radiation
  definition, 812
  incident, 819
  thermal, geometry, 817
Planets
  effective temperature, 112
  visual magnitudes, 112
Plastics
  absorption wavelengths, 326
  filter materials, 306-307
  Kel-F, 326
Platinum, 359
Plexiglass, 325
Pluto, 794
Points
  principal, optics, 371, 375
  singular
    linear systems, 672
    nonlinear systems, 704
    compensation methods, 723
Point sources, 51, 638
Point spread function, 615
Poisson distribution (noise), 585
Polar coordinates, 647
Polarization, filters, 286, 298
Polarizers, 298
Polyethylene, 325
Porro prism, 309
Potassium bromide
  Christiansen wavelength, 296
  prisms, 309
  reflection coefficient, 298
Potassium chloride, 298

Potassium iodide, 298
Power
  conversion factors, 879
  definition, 857
  dispersive, 339
  effective, radiometers, 763
  emissive, earth, 790
  gain, available, 588
  maximum transfer, 293
  resolving, 618
  responsivity, 761
  supply, 597
Power-density spectrum, 726
Prandtl number, 829
Preamplifiers
  available power gain, 588
  FET design, 608
  high-frequency compensation, 590
  low-noise design, 605
  mean-square noise voltage, 591
  noise factor, 589
Pressure
  conversion factors, 872
  definition, 856
  generator, Barnes nitrogen, 530
  regulator, 533
  total, atmospheric, 189
Principal point, optics, 371, 375
Principal rays, 399
Prisms
  Abbé, 308
  Amici, 308
  characteristics, 307
  constant-dispersing, 308
  deviating, 308
  dispersing, 308
  Dove, 308
  fore, 766
  interchangeable, 771
  inverting, 309
  Littrow, 308
  materials, 309
  monochromator, 766, 772
  Porro, 309
  position in infrared system, 736
  roof, 308
  rotation, 734
  spectroradiometric uses, 763
  total reflection, 309
  triple mirror, 309
  Wadsworth, 308
Probability
  of detection, 740
  gaussian, 740
Probes, deep space, 811
Projected solid angle, 24
Projection lenses, 426
Projection systems, 424
Propellant fuels, 62, 67
Proportionality condition, 614
Pulse response, detectors, 505
Pulse-position system, 746
Pupils, entrance and exit, 377-379, 616-621, 624, 636, 733
Pyromark standard black, 362

Quadrature method, gaussian, 631-633
Quantum efficiency, detective, 468
Quantum rates in blackbody radiation, 10
Quarter-wave optical thickness, 288
Quarter-wave stack, 288
Quartz
  Christiansen wavelengths, 296
  crystal, prisms, 309
  filter uses, substrate, 294
  glass, fused, 317
  iodine lamps, 40
  thermal radiation properties, 809
  windows, 39
Quasi-random model, 196-201

Radiance
  absorbed, 27
  background
    winter day, 164
    winter night, 165
  blackbody, 96
  brick wall, 146
  concrete, 146-156, 164
  definition, 784
  diurnal variations, 144, 149
  grass, 146, 157, 164
  ground, 143
  incident, 23, 27
  NBS standards, 41
  ocean, 169
  opaque body, 70
  reflected, 24, 27
  sky, 96-99, 143, 164
  snow, 146, 163-164
  terrain, *see* Terrain
  urban area, 143
Radiancy, 784
Radians, degree equivalents, 865
Radiant emissivity, *see* Emissivity
Radiant emittance, 23
Radiant energy reference level, 758
Radiant heat transfer, 787
Radiant intensity, 24
Radiation
  albedo, 790,812
  atmosphere, 737
  atmospheric temperature effects, 97
  background,
    discrimination, 737
    high-speed flight, 826
    minimization, 737, 739
  blackbody, 10
  boundary-layer effects, 846
  dosages, 360
  earth, 785
  equations and constants, 28
  geometry, 22
  heat transfer, 828
  heated air, 851

laws, 9-10
NBS standards, 38
planetary, 112, 812, 819
ratio of visible to infrared, 570
reference level, 759
rocket and jet propulsion systems, 68
reflected solar, 785, 790, 812, 814, 816
shock wave effects, 841
solar insolation, 96, 785, 812, 813
space, 820, 821
stellar, 107, 110
target, 737
thermal,
planetary, 817
properties of selected materials, 804
temperature of atmosphere, 97
total, of targets, 71
transducer, 458
ultraviolet, in space, 792
ultraviolet, effect on atmospheric ozone, 185
windows, 827, 833
Radiators, distributed, 23
Radiometers
alternating current, 760
basic design, 758
Cassegrain, 762
commercially available, 761
direct current, 759
effective power, 763
essential components, 758
normalization to the peak, 763
Pfund, 762
principal characteristics, 759
spectral bandwidth, 763
spectroradiometers, 763
Radiometric quantities, 4, 761, 784
Radiometric relations, 28
Radius of curvature, 840
Random Elsasser model, 196-201
Random process Fourier transfer, 649
Range
calculation, 747
conversion chart, 863
detection, 739
Rapid-scan monochromator, 773
Rapid-scan spectrometer, 773
Raster scan, 736
Rate error linear systems, 663
Rays
limitation by stops and aperture, 377
marginal, 399
marginal paraxial, 402
paraxial, 382, 399, 402
principal, 376-377, 399
principal paraxial, 402
skew trace, 396
Ray tracing
aberrations, 388
Coddington's equations, 392, 395
desk calculator, procedures, 390
electronic computers, 394
graphical, 398, 431
paraxial, 377, 390
precision, 388
series of elements, 375
Snell's law, 388
transfer function, 634
Rayleigh criterion, 412, 627
Rayleigh interferometers, 776
Rayleigh-Jeans law, 10, 20
Rayleigh scattering, 204, 208
Receiver-collector area, 224
Reciprocal dispersion, 403
Reciprocity, films and plates, 574
Reciprocity law, 27
Reciprocity-law failure, 574
Reciprocity theorem, Helmholtz, 24, 27
Recovery factor, 829
Rectangular aperture, 619
Rectangular aperture filter, 654
Reentry heating, 69
Reference blackbodies, 761
Reference radiation level, 759
Reflectance
aluminum foil, 76
asphalt, 83
coaltar pitch, 84
coating materials, 80, 354
definition, 784
diffuse, 70
directional, 23, 26
enamels, 80
fabrics, 76
high, zone filters, 280
masonry, 82
measured values, 74
nonequlibrium condition, 27
optical surface coatings, 351
ordnance materials, 74
paints, 80
partial, 25, 28
rubber, natural, 75
specular, 70
steel, 76
targets, 72
terrain features, 75, 85
total, 25, 28
Reflected solar radiation, *see* Albedo
Reflection, *see also* Reflectance
coefficient, 293, 297
distribution function, 24
filters, selective, 286, 297
internal
immersion lenses, 433
light pipes, 431
loss
different incidence angles, 350
tangent, 350
prism, total, 309
single layer, 348
single surface, 347
stop band, 290
water surface, 167
Reflectivity
angular dependence, 793
definition, 784

directional, 88-93, 793
interference filters, 292
maximum, filters, 288
partially transparent bodies, 349
reciprocity relation, 796
seawater, 166
selected materials, 796
terrain, 142
water, 167
wave slope, effect, 168
Reflectors, 437-440, 450
Refracting systems, 454
Refraction, 286, 298, 839
Refractive index, *see* Index of refraction
Relative noise temperature, 588
Relative stability (linear systems), 664
Relative visibility curve, films, 573
Relaxation methods of optical design, 409
Relay lenses, 426
multielement, 284
Relay systems, 424
Replica mirrors, 355-358
Representations-of-the-system method, 702
Residual aberrations, 386, 407
Resistance
detectors, 462
equivalent noise, 589, 593
Resistors, 590
Resolution
aberration limits, 414
criteria for point sources
Sparrow resolution, 628
Rayleigh resolution, 627
energy distribution, 410
limit, aerodynamic effects, 850
loss, due to turbulent boundary layer, 849
sine-wave, 628
spectral, in monochromators, 765
spurious, 629, 637
theoretical limit, 844
Resolving power, 618
Response
amplifiers, high-frequency, 595
apodization, 636
frequency
calculations of design, 680
computer calculation, 640
detectors, 504
FET amplifiers, 609
optical, 410
track loop, 746
impulse, 614
off-axis, 626
optical, 643
pulse, of detectors, 505
spectral, of detectors, 508
time, from phase plane, 710
transient, analysis, 667, 675, 702
Responsive area, detectors, 462
Responsive element, detector, 458
Responsive quantum efficiency, detectors, 468
Responsivity
blackbody, detectors, 464
power, 761
spectral, of detectors, 464
Reticles
circular sectored, 653
filters, circular sectored, 655
moving, 651
infinite checkerboard, 654, 655
infinite parallel spoke, 654, 685
square wave, checkerboard, 651
wagon-wheel, 653
Reynolds analogy, 829
Reynolds numbers, 828
Rhodium, 351, 809
Ribbon filament lamp, 58, 643
Rim-ray curve, 386-387
Rise time (linear systems), 662
RMS noise voltage, detectors, 464
RMS signal voltage, detectors, 463
Rocket and jet propulsion systems
afterburning heat release, 60
exhaust
composition, 62
emission bands, 64
equilibrium constant, 66
first-period length, 61
flow field, 61
in vacuum, 62
jet structure, 60
major emission bands, 67
molecular emission, 67
nozzles, 61, 62
particles, 67
stagnation temperature, 61
temperatures, 59
undisturbed cone, 62
velocity, 59
exit heating, 68
heat flow, 68
high-energy solid fuels, 67
liquid propellants, 64
mass flow, 59
propellants, 62, 67
radiation processes, 68
reentry heating, 69
shock-wave formation, 61
temperature, 68
thrust, 59
Rock salt prisms, 309
Rokide, 809
Roof prism (Amici), 308
Root locus methods,
linear system, 668, 683
sampled-data systems, 703
systems design, 722
Rosette scan, 736
Rotating prism, 734
Rotating wedges, 735
Rotation, rates of, 734, 736
Routh's stability test, 669
Rowland circle focal curve, 311
Rowland grating mounting, 312
Rubber, 75
Rubidium iodide, 296

s-plane
  contour (linear systems), 673
  transfer functions, 700
Sagittal coma, 382, 438-439
Sagittal focus, 384
Sample function (Fourier transforms), 649
Sampled-data systems
  analysis methods, 701
  absolute stability, 700-702
  definitions, 695
  design, 719
  transfer functions, 700
  types, 703
Sampler, 696
Sampling, 695
Sampling theorem, 696
Sapphire, 294
Satellites, 812
Saturation
  detectors, 834
  systems with, 713
Saturn, 794
Sawyer glass, 318
Scan patterns, 736
Scanning, 734, 747
Scanning-aperture space filters, 650
Scanning-field space filters, 652
Scattering, 96, 286
Scattering angle, 117
Scattering coefficients, 204-211
Schmidt system, 443, 453
Schur-Cohn stability criterion, 702
Scintillation
  See also Stars, stellar scintillation
  atmosphere, 210
  turbulent boundary-layer effect, 849
Sea, *see* Ocean
Sea water, 166
Search system design analysis, 737
Seasonal variations atmosphere, 186
Secondary maxima, diffraction pattern, 618
Seconds, radian equivalents, 865
Seebeck coefficient, 544, 548
Seidel aberrations, 382-390, 630
Selective absorption, 286
Selective radiators, 23
Selective reflection filters, 297
Selective refraction filters, 298
Selenium glass, 321
Semiconductors, 306-307, 324
Sensitivity
  analysis, 731, 744
  calculation, equations for, 731
  contours, detections, 510
  design analysis, 738
  front-end description, 747
  mapping, 731
  optical systems, 732
  parameters in system design, 740
  spectral, films, 574, 578
  target-detection, 731
  tracking systems, 744
Sensitometric characteristics, films, 573
Servo analysis, 745, 747
Servo-bandwidth-limited system, 750
Servo Corp. of America
  catadioptric systems, 285
  lenses, 282, 284
Servocon lenses, 282
Setting time (linear systems), 663
Seya-Namioka grating mounting, 312
Shape factor, 435
Sheet iron, 806
Shift, angle, interference filters, 290
Shimmer, *see* Atmosphere, scintillation; Stars, stellar scintillation
Shock effect on detectors, 461
Shock formation theory, 61
Shock waves
  aberrations, 842
  angle, 840
  effects, 839
    on focal length, 844
    on index of refraction, 844
    on infrared radiation, 841
    on resolution, 844
    on stagnation temperature, 83
  hemispherical, 846
  supersonic, 61
Short circuit impedance, 293
Short-wavepass interference filters, 290
Shot noise, 470, 585, 586, 589, 591, 592, 597
Sicon black, 47, 362
Signal flow diagrams, 664
Signal flow rules, 665
Signal generator, 589
Signal processing, 742
Signal-to-noise ratio, 594, 740, 748
Signal voltage, RMS, detectors, 463
Silicate glass, 321
Silicon
  detectors, 473
  effect of neutron bombardment on absorption coefficient, 358
  filter film, 295
  filter substrate, 294
  lenses, 282
  neutron bombardment, 358
  solar cell, 810
  transistor, low noise, 601
  transmission as semiconductor, 324
Silicon-oxygen group, 325
Silicon carbide, 359
Silicon:gold detectors, 496
Silicon monoxide
  emissivity, 359
  filter film, 295
  filter mirror coating, 352
  reflectance, 354
  transmittance, 356
Silicon:zinc detectors, 495
Silver, 351, 810
Silver chloride, 309
Silver paint, Codit, 90
Simple compensation networks, 718
Simple speed controllers, 686

Simpson's rule, 633
Simulator, blackbody, 32
Sine condition, 282, 389, 428
Sine waves
  modulation, 214
  resolution, 628
  spatial, 643
  target, 638
Single cell detector cooling systems, 557
Single-element lenses, 282
Single-element objectives, 434
Single-pass monochromator, 765
Single surface reflection and transmission, 347
Single-pulse system, 748
Single-stage multicouple cooling systems, 555
Single-stage thermoelectric cooling
    systems, 525, 554
Singular points, 672, 704, 723
Sintered metal filters, 533
Skew ray trace, 396
Sky, *see* Backgrounds, sky
Slant paths, atmospheric, 261-266
Slant range, 115
Slide rules, blackbody, 11
Slit
  aperture, 617-619, 630
  entrance, monochromator, 766
Slope (filters), 287
Smoke, 203, 356
Snells' law, 371, 388, 398
Snow, 142, 146, 163
Sodium aluminum fluoride, 356
Sodium bromide, 296
Sodium chloride, 295, 296, 298
Sodium fluoride, 298
Sodium iodide, 296
Solar absorptance, 792, 800
Solar activity, 101
Solar heat flux, 811
Solar radiation, *see* Radiation, solar
Solar spectrum measurements, 227-229
Solar system, 793
Solder, 810
Solid light pipes, 431
Solubility of optical materials, 333
Solution of aberration problems, 400
Soot, 359
Sources
  artificial, 32
  area, 223, 225
  blackbody (2400°K), 47
  blackbody simulator, 32
  carbon arc, 49
  cavity, 32, 51
  detector uses, 501
  field, 46
  Globar, 48
  laboratory, 46
  Lambertian, 22
  lamps, 49, 50
  low-temperature, 47
  mercury arc, 49, 50
  NBS, 38
  Nernst glower, 48
  Sicon-black enamel, 47
  Welsbach mantle, 48
  Zapon paint, 48
Space and space frequency differentiation, 647
Space charge-limited diodes, 586
Space environment, 792
  cooling systems, 561
Space-frequency shifting, 646
Space radiation, 820
Space scaling, 646
Space shifting, 646
Space technology terminology, 785
Spacecraft
  See also Rocket and jet propulsion systems
  components, temperature range, 822
  deep space probes, 811
  passive control, 822
  testing, 823
  thermal coatings, 787
  thermal design, 821
Sparrow method of calculating emissivity, 37
Sparrow resolution criteria, 628
Spatial domain, 623
Spatial filter analysis, 739
Spatial filtering, 737
Spatial frequencies, 626
Spatial frequency domain, 628
Spatial frequency filtering, 646
Spatial invariance, 621
Spatio-temporal filter, 649
Specific heat of optical materials, 333
Spectral bands and lines
  absorption, 238
  admittance (filters), 293
  carbon dioxide, 238, 239, 240, 242, 243
  carbon monoxide, atmospheric, 249
  equally intense distribution, 194
  exponential distribution of line intensities,
    194
  infinite distributed, 293
  infinite lumped-constant, 293
  intensity, 194
  Lorentz shape, 291
  methane, atmospheric, 250
  models, 192
  narrowness, 291
  nitrous oxide, atmospheric, 246
  terminated by z_R, 293
  transmission-line theory, filters, 293
  water vapor (atmospheric), 244
Spectral bandwidth, 763
Spectral classes of stars, 107
Spectral curve, Johnson solar, 789
Spectral D^*, 467
Spectral D^{**}, 467
Spectral detectivity, 466
Spectral distribution, 761
Spectral emissive power, earth, 790
Spectral emittance, 785
Spectral irradiance, *see* Irradiance
Spectral noise equivalent power, 466
Spectral purity, 765

Spectral radiance, *see* Radiance
Spectral reflectance, *see* Reflectance
Spectral response detectors, 464, 508
Spectral sensitivity, 574
Spectrometers
  commercially available, 767
  double-pass, 771
  flame temperature, 772
  grating, 766
  in-line, 774
  interferometer, 774
  Littrow mounting, 766
  Pfund grating, 767
  prism-grating double monochromator, 771
  rapid-scan, 773
Spectrophotometers
  direct-ratio, 776
  double-beam optical wall, 775
Spectroradiometers, 763
Spectroscopic energy, conversion factors, 878
Spectroscopic films, Kodak, 577, 572-573, 576-579
Spectrum analyzer, 726
Speed
  aperture, 379, 856, 869
  conversion factors, 856
  definition, 869
Speed controllers, simple control systems, 686
Spherical aberration, 382-388, 399-400, 414, 435-439, 446-447, 629-633
Spherical Fabry-Perot interferometer, 779
Spherical mirror, 437
Spherical reflector optical system, 437-439
Spherochromatism, 386
Spike, interference filters, 292
Spinning-mirror technique, 506
Spiral scan, 736
Spot diagram, 389
Spread function, 617
Spurious resolution, 629, 637
Square-band interference filter, 291
Square coefficient (absorptance), 196
Square-wave reticle filters, 651
Square root region, 191
Stability
  absolute
    analysis method, 669, 672, 674, 681, 685, 700
    linear systems, 664, 700-702
    nonlinear systems, 702
    sampled-data systems, 700
  asymptotic, 704
  criteria
    James-Weiss, 672
    Schur-Cohn, 702
  describing function analysis, 708
  global, 704
  Liaponov, 704, 710
  relative (linear systems), 664
  test, Routh's, 669
  transistors, 598
Stabilization
  bias, 598
  requirements of mapping systems, 755
  track, 750
Stack, quarter-wave, 288
Stagnation temperature, 61, 831
Stainless steel, 76, 801, 810
Standards
  atmospheric temperature profiles, 177
  carbon filament, 30
  length, 855
  NBS emissivity comparsion, 44
  NBS radiation, 38
  NBS spectral irradiance, 40
  NBS spectral radiance, 41
  noise temperature, 588
  Pyromark black, 362
Stanton number, 828
Stars
  apparent galactic concentration, 109
  background radiation, 107
  concentration by spectral class, 109
  effective temperature, 112
  numbers, by magnitude, 107
  spectral classes, percentages, 107
  spectral classification, 107
  spectral distribution of stellar radiation, 110
  spectral irradiance by visual magnitude, 114
  stellar scintillation
    aperture size, effect of, 218, 221
    crossover frequency, 215
    effects of upper air winds, 216
    zenith distance, 218, 222
  surface temperature, 107
  visual magnitudes, 112
Stationary systems, 614
Statistical model, 194-199
Steady-state behavior, 674, 681, 685
Steel, stainless, 76, 359, 810
Steepest descent design, optics, 408
Stefan-Boltzmann constant, 39
Stefan-Boltzmann law, 10
Stellite, 351
Step inputs, 672, 689, 693
Steradiancy, 784
Stirling cycle, 540
Stop band, 287, 290
Stop shift theory, 404, 440
Stratosphere
  aerosol content, 141
  clouds, 124
  dry, 181
  water vapor, 182, 183
  wet, 181
Straubel pupil function, 618
Strehl criterion, 621, 633, 636
Strip filament lamps, 50
Strong-line approximation of absorptance, 191, 200
Sub-auroral belts, 101
Substrates, filter, 287, 294
Sun, 794
  See also Insolation; headings beginning Solar
Superposition, linear systems, 614
Supersonic shock waves, 61

Surfaces
aplanatic, 427
aspheric, 385, 402
coatings, 352
irregularities, 25
Petzval, 384, 400, 439-440
targets, *see* Targets
thermopile, 40
Symmetrical principle, 406
Symmetry, axis of, 844
Synchronous detection, 610
Systems, *see* system names

T-12 (optical material), 326
Tabor, 810
Tangent, loss, 358
Tantalum, 803, 810
Target presence generation, 750
Target-detection sensitivity, 731
Targets, 58
See also Rocket and jet propulsion systems
blast furnaces, 69
concrete, 146
contrast, 71
directional reflectivity, 88-93, 793
greybody, 70
passive, temperatures of, 72
perfectly diffusing surface, 70
power plants, 69
radiance, 70
radiation design consideration, 737
reflectance, 70
supersonic shockwaves, 61
thermal emission, 70
total radiation, 71
vehicle exhausts, 69
Telescope
astronomical, 422
Cassegrain, 442
Galilian, 422
Gregorian, 442
Newtonian, 441
terrestrial, 422
Tellurium detectors, 483
Temperature
ambient, effects on radiance of
snow, 163
conversion factors, 873
definition, 856
detectors
background, 462
cycling tests, 461
noise, effect, 469
operating, 463
ranges of detector cooling system, 525
storage requirements, 461
effects on atmospheric scintillation, 223
effects on semiconductors, 324
gradients, 833
high-speed flight
rate of increase, 832
recovery, 832
wall, 829
local static, 831
low-, liquids, 520
lowest, for liquid gases, 521
maximum (recovery), 832
methane, solid, 562
noise, 588
optical materials
melting, 330
specific heat, 333
thermal conductivity, 331
thermal expansion, 332
profiles, 177-180
ranges, of spacecraft components, 822
sea-surface, distribution, 169
sea boundary layer structure, 170
stagnation, 61, 831
stellar, 107
terrain, 142
window, 833, 837
Terrain
airfield, 144
bare ground, 142-143
city, 144
concrete, *see* Concrete
diurnal variations in radiation, 143
dry sand, 144
emissivity, 75, 142
forest, 144
grass, 144, 146, 157, 164
moist sand, 144
ocean, 142
radiance, spectral, 142
radiation of selected backgrounds, 144
reflectance, 75, 85
reflectivity, 142
snow, 142, 146, 163, 164
spectra, 143
temperature, 142
urban area, 143
Terrestrial telescope, 422
Thallium
bromo-iodide, 309
bromide, 298
chloride, 297
iodide, 296, 298
Theoretical limit
detectivity, 512
resolution, 844
Thermal
coatings for spacecraft, 787
conductivity, 331
contact conductance, 786
detectors, 458, 459
emission, 70, 96
expansion of optical materials, 332
joint conduction, 786
loads, 827
measurements, 799
noise, 469, 585
radiation, *see* Radiation
Thermistors, 458, 498
Thermocouples, 458, 499
Thermoelectric cooling systems, 523, 544

Thermopiles, 40, 458
Thermopneumatic detectors, 458-459, 500
Thermovoltaic detectors, 459
Thevenin equivalent, 590
Thick lens, 375
Thickness, optical, 288, 294
Thin lens, 340, 372, 375, 404
Thin element optical systems, 436
Third-order aberrations, 382-390, 404, 414
Thomson relations, 548
Thorium fluoride, 295
Three-element achromat, 284
Threshold, operating, 740
Thrust, 59
Time
  behavior (linear systems), 672
  constant, 662
    detectors, 465, 504, 738
  conversion factors, 868
  definitions, 856
  delay, 662
  dwell, 730, 734, 739, 752
  integration, 745
  response, 663
    from phase-plane, 710
  rise, 662
  to peak, 663
Time-invariant systems, 614
Tin, 811
Titanium, 351, 810
Titanium dioxide, 295
Titanox-RA, 810
Torque
  conversion factors, 871
  definition, 856
Total absorptance, 28
Total emissivity, 28
Total emittance, 785
Total power, 437
Total radiation, 71, 821
Total reflection prism 309
Track accuracy, 746
Track stabilization, 750
Track-field size, 745
Track-loop
  definition, 744
  servo analysis, 747
  frequency response, 746
Track-while-scan system, 745
Tracking system, 744
  many tracking loops system, 692
Transconductance, dynamic, 586
Transducer, 458
Transfer, heat, *see* Heat transfer
Transfer efficiency, cooling systems, 531
Transfer functions
  See also Optical systems, transfer function
  closed loop, 722
  linear systems, 664
  s-plane, 700
  sampled-data systems, 700
Transform relationships, 696
Transformation analysis, 702
Transformers
  broadband, 604
  coupling, 604
  design, 604
Transient behavior, 674, 681, 685
Transient response analysis
  linear systems, 667, 675
  sampled-data systems, 702
Transistors, 586, 597-602, 604, 606
Transmission
  atmospheric, 737
  efficiency, light pipes, 431
  Lyot filter, 299
  semiconductors, 307
  through the atmosphere, 252
  through clouds, 258
  window, at elevated temperatures, 833
Transmission-line theory, 293
Transmission region of optical materials, 326
Transmissivity
  calculation of atmospheric slant paths, 266
  partially transparent bodies, 349
  sea water, 166
Transmittance
  antireflection coatings, 356
  cooled optical materials, 359-361
  definition, 784
  filters (photographic), 574, 580
  glass
    calcium aluminate, 323
    Corning No. 18655, 318
    Germanate, 322
    high-silica, 318
  interferometers, 780
  optical materials, 364
  peak, 287
  polyethylene, 325
  semiconductor materials, 324
  single layer, 348
  single surface, 347
  substrate, 287
Transverse aberrations, 381, 385, 390, 399
Trap, cold, 533
Triode, 589, 592
Triple mirror prism, 309
Triple point, 520
Tropopause, 122-123
Tubes, vacuum, 592-595
Tungsten, 803, 811
Tungsten filament lamp, 40, 49, 643
Tungsten oxide, 356
Turbine expander cooling systems, 544, 545
Turbulent flow
  aerodynamic effects, 848
  recovery factor, 829
  Stanton number, 831
  transition from laminar flow, 828
Twilight airglow, 106
Twinkle, *see* Atmosphere, scintillation
Twyman-Green interferometer, 777
Type 0 linear system, 686, 703
Type 1 linear system, 688, 703
Type 2 linear system, 692, 703

Ultraviolet absorptance, 792
Ultraviolet radiation in space, 792
Upper air winds, 216
Uranus, 794

V number, Abbé, 403, 409
V value, 435
Vacuum environment specifications, detectors, 461
Vacuum tubes, 592, 593, 595
  amplifiers, 592
Velocity
  conversion factors, 869
  definition, 856
Venus, 794
Vertical blackbody, 45
Very-high-impedance amplifiers, 596
Vibration requirements, detectors, 461
View angles, 115
Visibility; relative, 573
Vignetting, 378, 429
Vinyl, 76
Voltage
  output noise, 599
  RMS signal, detectors, 463, 464
Volume
  conversion factors, 864
  definition, 855
Vycor glass, 294, 317

Wadsworth grating mounting, 312
Wadsworth prism, 308
Water, 167, 362
  sea, 167
Water vapor, 96, 178, 181, 244
Wave, symmetrical triangle; 651
Wave number, 288
Wave slope, 168
Wave-front aberration, 616
Wavelength
  detectors
    cutoff, 468
    peak, 467
  filters, 287
  operation, 738
Weak-line approximation, 197-198
Wedges, rotating, 735
Weighting function (linear systems), 663, 672
Welsbach mantle, 48
White enamel, 359
Wind, 216, 223
Wien law, 10, 20
Wiener spectrum, 649, 651
Windows
  aberrations, 826
  detectors, 359
  ease of fabrication, 827
  emissivity, 827
  entrance and exit, 378
  fluorite, 39
  heating of, 827
  hot-, problem alleviation, 836
  materials
    characteristics, 833
    emissivity, 835
    magnesium fluoride, 836
  parameters, 836
  quartz, 39
  radiating, effects, 833
  radiation, 827
    at elevated temperature, 833
  requirements, 826
  temperature rise, 837
  transmission at elevated temperature, 833
Wool, 76
Work
  conversion factors, 875
  definition, 857

Xenon, 179

Young's modulus, 333

z-plane, transfer functions, 700
z-transform, 698
Zapon paint, 48
Zenith, 96, 99, 105, 115, 143
Zero-order circuits, 696
Zinc
  Ge-Si:Zn detectors, 495
  Ge:Zn detectors, 493, 495
Zinc sulfide, 295, 297, 356
Zinc oxide paint, 86
Zirconia, 359
Zirconium, 51
Zirconium dioxide, 295
Zonal aberrations, 386
Zone, auroral, 101
Zone, high-reflectance, 288-290

☆ U. S. GOVERNMENT PRINTING OFFICE : 1966 O - 236-960

LIBRARY OF CONGRESS

0 011 521 059 2